



SEQUENCE LISTING

<110> Barnett, Susan
Zur Megede, Jan

<120> POLYNUCLEOTIDES ENCODING ANTIGENIC HIV TYPE C
POLYPEPTIDES, POLYPEPTIDES AND USES THEREOF

<130> PP01631.101

<140> 09/475,704

<141> 1999-12-30

<150> 09/610,313

<151> 2000-07-05

<160> 46

<170> PatentIn Ver. 2.0

<210> 1

<211> 60

<212> DNA

<213> Human immunodeficiency virus

<400> 1

gacatcaagc agggccccaaggagcccttc cgcgactacg tggaccgctt cttcaagacc 60

<210> 2

<211> 60

<212> DNA

<213> Human immunodeficiency virus

<400> 2

gacatccgcc agggccccaaggagcccttc cgcgactacg tggaccgctt cttcaagacc 60

<210> 3

<211> 1479

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic Gag
of HIV strain AF110965

<400> 3

atgggcgccc gcgccagcat cctgcgcggc ggcaagctgg acgcctggga ggcgatccgc 60
ctgcgccccg gcggaagaa gtgctacatg atgaagcacc tgggtgtggc cagccgcgag 120
ctggagaagt tcgccctgaa ccccgccctg ctggagacca gcgagggtg caagcagatc 180
atccgccagc tgcacccgcg cctgcagacc ggcagcgagg agctgaagag cctgttcaac 240
accgtggcca ccctgtactg cgtgcacgag aagatcgagg tccgcgacac caaggaggcc 300
ctggacaaga tcgaggagga gcagaacaag tgccagcaga agatccagca ggccgaggcc 360
gccgacaagg gcaaggtgag ccagaactac cccatcgtgc agaacctgca gggccagatg 420
gtgcaccagg ccatcagccc ccgcaccctg aacgcctggg tgaaggtgat cgaggagaag 480
gccttcagcc ccgaggtgat ccccatgttc accgcctga gcgaggggcg cccccccag 540
gacctgaaca cgatgttgaa caccgtgggc ggccaccagg ccgccatgca gatgctgaag 600
gacacatca acgaggaggc cgccgagtgg gaccgcgtgc accccgtgca cgccggcccc 660

atcgcccccg	gccagatgcg	cgagccccgc	ggcagcgaca	tcgccggcac	caccagcacc	720
ctgcaggagc	agatcgccctg	gatgaccagc	aaccccccca	tccccgtggg	cgacatctac	780
aagcgggtga	tcctcctggg	cctgaacaag	atcgtgcgga	tgtacagccc	cgtgagcatc	840
ctggacatca	agcagggccc	caaggagccc	ttccgcgact	acgtggaccg	cttcttcaag	900
accctgcgcg	ccgagcagag	caccagggag	gtgaagaact	ggatgaccga	caccctgctg	960
gtgcagaacg	ccaaccccga	ctgcaagacc	atcctgcgcg	ctctcggccc	cggcgccagc	1020
ctggaggaga	tgatgaccgc	ctgccagggc	gtgggcgggc	ccagccacaa	ggccccgctg	1080
ctggccgagg	cgatgagcca	ggccaacacc	agcgtgatga	tgcagaagag	caacttcaag	1140
ggcccccggc	gcctcgtcaa	gtgcttcaac	tgcggcaagg	agggccacat	cgccccgcaac	1200
tgccgcgccc	cccgaagaa	gggctgctgg	aagtgcggca	aggagggcca	ccagatgaag	1260
gactgcaccg	agcggcaggc	caacttctctg	ggcaagatct	ggcccagcca	caagggccgc	1320
cccggcaact	tcctgcagag	ccgccccgag	cccaccgccc	cccccgccga	gagcttccgc	1380
ttcgaggaga	ccacccccgg	ccagaagcag	gagagcaagg	accgcgagac	cctgaccagc	1440
ctgaagagcc	tggttcggcaa	cgacccccctg	agccagtaa			1479

<210> 4

<211> 1509

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic Gag
of HIV strain AF110967

<400> 4

atgggcgccc	gcgccagcat	cctgcgcggc	gagaagctgg	acaagtggga	gaagatccgc	60
ctgcgccccg	gcggaagaa	gcactacatg	ctgaagcacc	tggtgtgggc	cagccgcgag	120
ctggagggct	tcgcctgaa	ccccggcctg	ctggagaccg	ccgagggctg	caagcagatc	180
atgaagcagc	tgagcccgcc	cctgcagacc	ggcaccgagg	agctgcgcag	cctgtacaac	240
accgtggcca	ccctgtactg	cgtgcacggc	ggcatcgagg	tccgcgacac	caaggaggcc	300
ctggacaaga	tcgaggagga	gcagaacaag	tcccagcaga	agaccagca	ggccaaggag	360
gcccagcgga	aggtgagcca	gaactacccc	atcgtgcaga	acctgcaggg	ccagatgggtg	420
caccaggcca	tcagcccccg	caccctgaac	gcctgggtga	aggtgatcga	ggagaaggcc	480
ttcagccccg	aggtgatccc	catgttcacc	gccctgagcg	agggcgccac	ccccaggac	540
ctgaacacga	tggtgaacac	cgtgggcggc	caccaggccg	ccatgcagat	gctgaaggac	600
accatcaacg	aggaggccgc	cgagtgggac	cgctgcacc	ccgtgcaggc	cgccccctg	660
gcccccgccc	agatgcgcga	ccccgcggc	agcgacatcg	ccggcgccac	cagcaccctg	720
caggagcaga	tcgcctggat	gaccagcaac	ccccccgtgc	ccgtgggcga	catctacaag	780
cggtggatca	tcctgggcct	gaacaagatc	gtgcggatgt	acagccccgt	gagcatcctg	840
gacatccgcc	agggcccca	ggagcccttc	cgcgactacg	tggaccgctt	cttcaagacc	900
ctgcgcgccc	agcaggccac	ccaggacgtg	aagaactgga	tgaccgagac	cctgctggtg	960
cagaacgcca	accccgactg	caagaccatc	ctgcgcgctc	tcggccccgg	cgccaccctg	1020
gaggagatga	tgaccgcctg	ccagggcgtg	ggcggccccg	gccacaaggc	ccgcgtgctg	1080
gccgagggca	tgagccaggc	caacagcgtg	aacatcatga	tgcagaagag	caacttcaag	1140
ggcccccggc	gcaacgtcaa	gtgcttcaac	tgcggcaagg	agggccacat	cgccaagaac	1200
tgccgcgccc	cccgaagaa	gggctgctgg	aagtgcggca	aggagggcca	ccagatgaag	1260
gactgcaccg	agcggcaggc	caacttctctg	ggcaagatct	ggcccagcca	caagggccgc	1320
cccggcaact	tcctgcagaa	ccgcagcgag	cccgcgccc	ccaccgtgcc	caccgcccc	1380
cccgccgaga	gcttcgctt	cgaggagacc	acccccgccc	ccaagcagga	gccaaggac	1440
cgcgagccct	accgcgagcc	cctgaccgcc	ctgcgcagcc	tggttcggcag	cggccccctg	1500
agccagtaa						1509

<210> 5

<211> 141

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Env common
region of HIV strain AF110968

<400> 5

```
accatcacca tcacctgccg catcaagcag atcatcaaca tgtggcagaa ggtgggccgc 60
gccatgtacg cccccccat cgccggcaac ctgacctgcg agagcaacat caccggcctg 120
ctgctgaccc gcgacggcgg c                                     141
```

<210> 6

<211> 1431

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
gpl20 coding region of HIV strain AF110968

<400> 6

```
agcgtggtgg gcaacctgtg ggtgaccgtg tactacggcg tgcccgtgtg gaaggaggcc 60
aagaccaccc tgttctgcac cagcgacgcc aaggcctacg agaccgaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gccaccgcac cccaaccccc aggagatcgt gctggagaac 180
gtgaccgaga acttcaacat gtggaagaac gacatgggtg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc tgcgtgaagc tgacccccct gtgctgacc 300
ctgaagtgcc gcaacgtgaa cgccaccaac aacatcaaca gcatgatcga caacagcaac 360
aagggcgaga tgaagaactg cagcttcaac gtgaccaccg agctgcgcga ccgcaagcag 420
gaggtgcacg ccctgttcta ccgcctggac gtggtgcccc tgcagggcaa caacagcaac 480
gagtaccgcc tgatcaactg caacaccagc gccatcacc aggcctgcc caaggtgagc 540
ttcgacccc tccccatcca ctactgcacc ccgcggcgt acgcatcct gaagtgaac 600
aaccagacct tcaacggcac cggcccctgc aacaacgtga gcagcgtgca gtgcgccac 660
ggcatcaagc ccgtggtgag caccagctg ctgctgaacg gcagcctggc caagggcgag 720
atcatcatcc gcagcgagaa cctggccaac aacgccaaga tcatcatcgt gcagctgaac 780
aagcccgtga agatcgtgtg cgtgcgcccc aacaacaaca ccgcaagag cgtgcgcac 840
ggccccggcc agacattcta cgccaccggc gagatcatcg gcgacatccg ccaggcctac 900
tgcatcatca acaagaccga gtggaacagc accctgcagg gcgtgagcaa gaagctggag 960
gagcacttca gcaagaaggc catcaagtgc gagcccagca gcggcggcga cctggagatc 1020
accaccaca gtttcaactg ccgcggcgag ttcttctact gcgacaccag ccagctgttc 1080
aacagcacct acagccccag cttcaacggc accgagaaca agctgaacgg caccatcacc 1140
atcacctgcc gcatcaagca gatcatcaac atgtggcaga aggtgggccg cgccatgtac 1200
gcccccccca tcgccggcaa cctgacctgc gagagcaaca tcaccggcct gctgctgacc 1260
cgcgacggcg gcaagaccgg cccaacgcac accgagatct tccgccccgg cggcggcgac 1320
atgcgcgaca actggcgcaa cgagctgtac aagtacaagg tggaggagat caagcccctg 1380
ggcgtggccc ccaccgaggc caagcgccgc gtggtggagc gcgagaagcg c                                     1431
```

<210> 7

<211> 1944

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
gpl40 coding region of HIV strain AF110968

<400> 7

```
agcgtggtgg gcaacctgtg ggtgaccgtg tactacggcg tgcccgtgtg gaaggaggcc 60
aagaccaccc tgttctgcac cagcgacgcc aaggcctacg agaccgaggt gcacaacgtg 120
```

tggggcacc	acgcctg	gcccaccg	cccaacccc	aggagatc	gctggaga	180
gtgaccgag	acttcaac	gtggaaga	gacatggt	accagatg	cgaggacat	240
atcagcctg	gggaccag	cctgaagcc	tgcgtgaag	tgaccccc	gtgctgacc	300
ctgaagtgc	gcaacgtg	cgccacca	aacatcaac	gcatgatcg	caacagca	360
aagggcgag	tgaagaact	cagcttca	gtgaccacc	agctgcgcg	ccgcaagc	420
gaggtgcac	ccctgttct	ccgcctgg	gtggtgccc	tgaggggca	caacagca	480
gagtaccgc	tgatcaact	caacaccag	gccatcccc	aggcctgcc	caaggtgag	540
ttcgacccc	tccccatcc	ctactgcac	ccgcgcgg	acgccatcc	gaagtgaac	600
aaccagacct	tcaacggc	cgccccctg	aacaacgtg	gcagcgtgc	gtgcgccc	660
ggcatcaag	ccgtggtg	caccagctg	ctgctgaac	gcagcctgg	caagggcg	720
atcatcatc	gcagcgaga	cctggcca	aacgccaag	tcatcatcg	gcagctga	780
aagcccgtg	agatcgtgt	cgtgcgccc	aacaacaac	cccgaagag	cgtgcgcac	840
ggccccggc	agaccttct	cgccaccgg	gagatcatc	gcgacatcc	ccaggccta	900
tgcatcatc	acaagaccg	gtggaacag	accctgcag	gcgtgagca	gaagctgg	960
gagcacttc	gcaagaagg	catcaagtt	gagcccagc	gcggcggcg	cctggagat	1020
accaccaca	gcttcaact	ccgcggcg	ttcttctac	gcgacacc	ccagctgtt	1080
aacagcacct	acagcccc	cttcaacgg	accgagaac	agctgaacg	caccatcac	1140
atcacctgc	gcatcaagc	gatcatca	atgtggcag	aggtggg	cgccatgt	1200
gcccccccc	tcgcccgg	cctgacctg	gagagcaac	tcaccggc	gctgctgac	1260
cgcgacggc	gcaagaccg	ccccaacg	accgagatc	tccgcccc	cggcggcg	1320
atgcgcgac	actggcgca	cgagctgt	aagtacaag	tggtgggat	caagcccct	1380
ggcgtggcc	ccaccgagg	caagcgccg	gtggtggag	gcgagaagc	cgccgtggg	1440
atcggcgcg	tgcttctgg	cttctctgg	gccgcccgc	gcaccatgg	cgccgccag	1500
atcacctga	ccgtgcagg	ccgcctgct	ctgagcggc	tcgtgcagc	gcagaaca	1560
ctgctgcgc	ccatcgagg	ccagcagca	ctgctgcag	tgaccgtgt	gggcatca	1620
cagctgcag	ccgcctacc	ggcctggag	cgctacctg	aggaccagc	gctgctggg	1680
atctggggc	gcagcggca	gctgatctg	accaccgcc	tgccctgga	cagcagctg	1740
agcaaccgc	gccacgacg	gatctggg	aacatgacct	ggatgcagt	ggaccgcg	1800
atcaacaact	acaccgac	catctaccg	ctgctggag	agagccaga	ccagcagg	1860
aagaacgag	aggacctgt	ggccctgg	agctggcag	acctgtgga	ctggttcag	1920
atcaccaact	ggctgtgga	catc				1944

<210> 8

<211> 2466

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
gp160 coding region of HIV strain AF110968

<400> 8

agcgtggtg	gcaacctgt	ggtgaccgt	tactacggc	tgcccgtgt	gaaggaggc	60
aagaccacc	tgttctgc	cagcgacgc	aaggcctac	agaccgagg	gcacaacgt	120
tggggcacc	acgcctg	gcccaccg	cccaacccc	aggagatc	gctggaga	180
gtgaccgag	acttcaac	gtggaaga	gacatggt	accagatg	cgaggacat	240
atcagcctg	gggaccag	cctgaagcc	tgcgtgaag	tgaccccc	gtgctgacc	300
ctgaagtgc	gcaacgtg	cgccacca	aacatcaac	gcatgatcg	caacagca	360
aagggcgag	tgaagaact	cagcttca	gtgaccacc	agctgcgcg	ccgcaagc	420
gaggtgcac	ccctgttct	ccgcctgg	gtggtgccc	tgaggggca	caacagca	480
gagtaccgc	tgatcaact	caacaccag	gccatcccc	aggcctgcc	caaggtgag	540
ttcgacccc	tccccatcc	ctactgcac	ccgcgcgg	acgccatcc	gaagtgaac	600
aaccagacct	tcaacggc	cgccccctg	aacaacgtg	gcagcgtgc	gtgcgccc	660
ggcatcaag	ccgtggtg	caccagctg	ctgctgaac	gcagcctgg	caagggcg	720
atcatcatc	gcagcgaga	cctggcca	aacgccaag	tcatcatcg	gcagctga	780
aagcccgtg	agatcgtgt	cgtgcgccc	aacaacaac	cccgaagag	cgtgcgcac	840
ggccccggc	agaccttct	cgccaccgg	gagatcatc	gcgacatcc	ccaggccta	900

tgcatcatca	acaagaccga	gtggaacagc	accctgcagg	gcgtgagcaa	gaagctggag	960
gagcacttca	gcaagaaggc	catcaagttc	gagcccagca	gcggcggcga	cctggagatc	1020
accacccaca	gcttcaactg	ccgcggcgag	ttcttctact	gcgacaccag	ccagctgttc	1080
aacagcacct	acagccccag	cttcaacggc	accgagaaca	agctgaacgg	caccatcacc	1140
atcacctgcc	gcatcaagca	gatcatcaac	atgtggcaga	aggtgggccc	cgccatgtac	1200
gcccccccca	tcgccggcaa	cctgacctgc	gagagcaaca	tcaccggcct	gctgctgacc	1260
cgcgacggcg	gcaagaccgg	ccccaacgac	accgagatct	tccgccccgg	cggcggcgac	1320
atgcgcgaca	actggcgcaa	cgagctgtac	aagtacaagg	tggaggagat	caagcccctg	1380
ggcgtggccc	ccaccgaggc	caagcgccgc	gtggtggagc	gcgagaagcg	cgccgtgggc	1440
atcggcgccg	tgttcctggg	cttcctgggc	gccgcgggca	gcacatggg	cgccgccagc	1500
atcacctga	ccgtgcaggc	ccgcctgctg	ctgagcggca	tcgtgcagca	gcagaacaac	1560
ctgctgcgcg	ccatcgaggc	ccagcagcac	ctgctgcagc	tgaccgtgtg	gggcatcaag	1620
cagctgcaga	cccgcctcct	ggcctgggag	cgctacctga	aggaccagca	gctgctgggc	1680
atctggggct	gcagcggcaa	gctgatctgc	accaccggcg	tgccctggaa	cagcagctgg	1740
agcaaccgca	gccacgacga	gatctgggac	aacatgacct	ggatgcagtg	ggaccgcgag	1800
atcaacaact	acaccgacac	catctaccgc	ctgctggagg	agagccagaa	ccagcaggag	1860
aagaacgaga	aggacctgct	ggccctggac	agctggcaga	acctgtggaa	ctggttcagc	1920
atcaccaact	ggctgtggta	catcaagatc	ttcatcatga	tcgtgggcgg	cctgatcggc	1980
ctgcgcatca	tcttcgccgt	gctgagcatc	gtgaaccgcg	tgcgccaggg	ctacagcccc	2040
ctgcccttcc	agacctgac	ccccaacccc	cgcgagcccc	accgcctggg	ccgcctcgag	2100
gaggagggcg	gcgagcagga	ccgcggccgc	agcatccgcc	tggtagcgcg	cttcctggcc	2160
ctggcctggg	acgacctgcg	cagcctgtgc	ctgttcagct	accaccgcct	gcgcgacttc	2220
atcctgatcg	ccgcccgcgt	gctggagctg	ctgggccagc	gcggctggga	ggccctgaag	2280
tacctgggca	gcctggtgca	gtactggggc	ctggagctga	agaagagcgc	catcagcctg	2340
ctggacacca	tcgccatcgc	cgtggccgag	ggcaccgacc	gcatcatcga	gttcatccag	2400
cgcatctgcc	gcgccatccg	caacatcccc	cgccgcatcc	gccagggcct	cgaggccgcc	2460
ctgcag						2466

<210> 9

<211> 2547

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
signal sequence and gpl60 coding region of HIV
strain AF110968

<400> 9

atgcgcgtga	tgggcatcct	gaagaactac	cagcagtggt	ggatgtgggg	catcctgggc	60
ttctggatgc	tgatcatcag	cagcgtgggt	ggcaacctgt	gggtgaccgt	gtactacggc	120
gtgcccgtgt	ggaaggaggc	caagaccacc	ctgttctgca	ccagcgacgc	caaggcctac	180
gagaccgagg	tgacaacgt	gtgggccacc	cacgcctgcg	tgcccaccga	ccccaacccc	240
caggagatcg	tgctggagaa	cgtgaccgag	aacttcaaca	tgtggaagaa	cgacatggtg	300
gaccagatgc	acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgcgtgaag	360
ctgaccccc	tgtgcgtgac	cctgaagtgc	cgcaacgtga	acgccaccaa	caacatcaac	420
agcatgatcg	acaacagcaa	caagggcgag	atgaagaact	gcagcttcaa	cgtgaccacc	480
gagctgcgcg	accgcaagca	ggaggtgcac	gccctgttct	accgcctgga	cgtggtgccc	540
ctgcagggca	acaacagcaa	cgagtaccgc	ctgatcaact	gcaacaccag	cgccatcacc	600
caggcctgcc	ccaaggtgag	cttcgacccc	atccccatcc	actactgcac	ccccgcccgc	660
tacgccatcc	tgaagtgcaa	caaccagacc	ttcaacggca	ccggccccctg	caacaacgtg	720
agcagcgtgc	agtgcgcccc	cgcatcaag	cccgtggtga	gcacccagct	gctgctgaac	780
ggcagcctgg	ccaagggcga	gatcatcatc	cgcagcgaga	acctggccaa	caacgccaaag	840
atcatcatcg	tgagctgaa	caagcccgtg	aagatcgtgt	gcgtgcgccc	caacaacaac	900
acccgcaaga	gcgtgcgcat	cggccccggc	cagaccttct	acgccaccgg	cgagatcatc	960
ggcgacatcc	gccaggccta	ctgcatcatc	aacaagaccg	agtggaacag	caccctgcag	1020
ggcgtgagca	agaagctgga	ggagcacttc	agcaagaagg	ccatcaagtt	cgagcccagc	1080

agcggcgggcg	acctggagat	caccacccac	agcttcaact	gccgcggcga	gttctttctac	1140
tgcgacacca	gccagctgtt	caacagcacc	tacagcccca	gcttcaacgg	caccgagaac	1200
aagctgaacg	gcaccatcac	catcacctgc	cgcacaaagc	agatcatcaa	catgtggcag	1260
aaggtggggc	gcgccatgta	cgcccccccc	atcgccggca	acctgacctg	cgagagcaac	1320
atcaccggcc	tgctgctgac	ccgcgacggc	ggcaagaccg	gccccaacga	caccgagatc	1380
ttccgccccg	gcggcggcga	catgcgcgac	aactggcgca	acgagctgta	caagtacaag	1440
gtggtggaga	tcaagcccct	gggcgtggcc	cccaccgagg	ccaagcgccg	cgtggtggag	1500
cgcgagaagc	gcgccgtggg	catcggcgcc	gtgttcctgg	gcttcctggg	cgccgccggc	1560
agcaccatgg	gcgccgccag	catcaccttg	accgtgcagg	cccgcctgct	gctgagcggc	1620
atcgtgcagc	agcagaacaa	cctgctgcgc	gccatcgagg	cccagcagca	cctgctgcag	1680
ctgaccgtgt	ggggcatcaa	gcagctgcag	acccgcatcc	tggccgtgga	gcgctacctg	1740
aaggaccagc	agctgctggg	catctggggc	tgcagcggca	agctgatctg	caccaccgcc	1800
gtgccctgga	acagcagctg	gagcaaccgc	agccacgacg	agatctggga	caacatgacc	1860
tggatgcagt	gggaccgcga	gatcaacaac	tacaccgaca	ccatctaccg	cctgctggag	1920
gagagccaga	accagcagga	gaagaacgag	aaggacctgc	tggccctgga	cagctggcag	1980
aacctgtgga	actggttcag	catcaccaac	tggctgtggt	acatcaagat	cttcatcatg	2040
atcgtggggc	gcctgatcgg	cctgcgcatc	atcttcgccg	tgctgagcat	cgtgaaccgc	2100
gtgcgccagg	gctacagccc	cctgcccttc	cagaccctga	cccccaacc	ccgcgagccc	2160
gaccgcctgg	gccgcacga	ggaggagggc	ggcgagcagg	accgcggccg	cagcatccgc	2220
ctggtgagcg	gcttcctggc	cctggcctgg	gacgacctgc	gcagcctgtg	cctgttcagc	2280
taccaccgcc	tgcgcgactt	catcctgac	gccgcccgcg	tgctggagct	gctgggccag	2340
cgcggtctgg	aggccctgaa	gtacctgggc	agcctggtgc	agtactgggg	cctggagctg	2400
aagaagagcg	ccatcagcct	gctggacacc	atcgccatcg	ccgtggccga	gggcaccgac	2460
cgcatcatcg	agttcatcca	gcgcatctgc	cgcgccatcc	gcaacatccc	ccgccgcgac	2520
cgccagggct	tcgaggccgc	cctgcag				2547

<210> 10

<211> 1035

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic a
gp41 coding region of HIV strain AF110968

<400> 10

gccgtgggca	tcggcgccgt	gttcctgggc	ttcctggggc	ccgccggcag	caccatgggc	60
gccgccagca	tcaccctgac	cgtgcaggcc	cgctgctgc	tgagcggcat	cgtgcagcag	120
cagaacaacc	tgctgcgcgc	catcgaggcc	cagcagcacc	tgctgcagct	gaccgtgtgg	180
ggcatcaagc	agctgcagac	ccgcacccct	gccgtggagc	gctacctgaa	ggaccagcag	240
ctgctgggca	tctggggctg	cagcggcaag	ctgatctgca	ccaccgccgt	gccctggaac	300
agcagctgga	gcaaccgcag	ccacgacgag	atctgggaca	acatgacctg	gatgcagtgg	360
gaccgcgaga	tcaacaacta	caccgacacc	atctaccgcc	tgctggagga	gagccagaac	420
cagcaggaga	agaacgagaa	ggacctgctg	gccctggaca	gctggcagaa	cctgtggaac	480
tggttcagca	tcaccaactg	gctgtggtac	atcaagatct	tcatcatgat	cgtgggcggc	540
ctgatcggcc	tgcgcatcat	cttcgccgtg	ctgagcatcg	tgaaccgcgt	gcgccagggc	600
tacagcccc	tgcccttcca	gacctgacc	cccaaccccc	gcgagcccga	ccgcctgggc	660
cgcacgcagg	aggagggcgg	cgagcaggac	cgcgcccgca	gcatccgcct	ggtgagcggc	720
ttcctggccc	tggcctggga	cgacctgcgc	agcctgtgcc	tggtcagcta	ccaccgcctg	780
cgcgacttca	tccctgatcgc	cgcccgcgtg	ctggagctgc	tgggcccagc	cggctgggag	840
gccctgaagt	acctgggcag	cctggtgcag	tactggggcc	tggagctgaa	gaagagcgcc	900
atcagcctgc	tggacaccat	cgccatcgcc	gtggccgagg	gcaccgaccg	catcatcgag	960
ttcatccagc	gcatctgccg	cgccatccgc	aacatccccc	gccgcatccg	ccagggcttc	1020
gaggccgccc	tgag					1035

<210> 11

<211> 144

<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic Env
common region of HIV strain AF110975

<400> 11

```
agcatcatca ccctgccctg ccgcatcaag cagatcatcg acatgtggca gaaggtgggc 60
cgcgccatct acgccccccc catcgagggc aacatcacct gcagcagcag catcaccggc 120
ctgctgctgg cccgcgacgg cggc                                     144
```

<210> 12

<211> 1437

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
gp120 coding region of HIV strain AF110975

<400> 12

```
agcggcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60
agcaccaccc tgttctgcgc cagcgacgcc aaggcctacg agaaggaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gcccaccgac cccaaccccc aggagatcga gctggacaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc cgcgatgaagc tgacccccct gtgctgaccc 300
ctgaagtgca ccaactacag caccaactac agcaacacca tgaacgccac cagctacaac 360
aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccga gctgcgcgac 420
aagaagcagc aggtgtacgc cctgttctac aagctggaga tcgtgcccct gaacagcaac 480
agcagcgagt accgcctgat caactgcaac accagcgcca tcaccaggc ctgccccaaag 540
gtgagcttgc accccatccc catccactac tgcgcccccg ccggctacgc catcctgaag 600
tgcaagaaca acaccagcaa cggcaccggc ccctgccaga acgtgagcac cgtgcagtgc 660
acccacggca tcaagcccggt ggtgagcacc cccctgctgc tgaacggcag cctggccgag 720
ggcggcgaga tcatcatccg cagcaagaac ctgagcaaca acgcctacac catcatcgtg 780
cacctgaacg acagcgtgga gatcgtgtgc acccgcccc aacaacaac cgcgaagggc 840
atccgcatcg gccccggcca gaccttctac gccaccgaga acatcatcgg cgacatccgc 900
caggcccact gcaacatcag cgccggcgag tggacaagg ccgtgcagcg cgtgagcgcc 960
aagctgcgcg agcatttccc caacaagacc atcgagttcc agcccagcag cggcggcgac 1020
ctggagatca ccaccacag cttcaactgc cgcggcgagt tcttctactg caacaccagc 1080
aagctgttca acagcagcta caacggcacc agctaccgcg gcaccgagag caacagcagc 1140
atcatcacc tgccttgccg catcaagcag atcatcgaca tgtggcagaa ggtgggcccgc 1200
gccatctacg ccccccccat cgagggcaac atcacctgca gcagcagcat caccggcctg 1260
ctgctggccc gcgacggcgg cctggacaac atcaccaccg agatcttccg cccccagggc 1320
ggcgacatga aggacaactg gcgcaacgag ctgtacaagt acaaggtggt ggagatcaag 1380
ccctggggcg tggccccac cgaggccaag cgccgcgtgg tggagcgaga gaagcgc 1437
```

<210> 13

<211> 1950

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
gp140 coding region of HIV strain AF110975

<400> 13

```

agcggcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60
agcaccaccc tgttctgcgc cagcgacgcc aaggcctacg agaaggaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gccaccgac cccaaccccc aggagatcga gctggacaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc cgcgtaagc tgacccccct gtgctgacc 300
ctgaagtgca ccaactacag caccaactac agcaacacca tgaacgccac cagctacaac 360
aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccga gctgcgcgac 420
aagaagcagc aggtgtacgc cctgttctac aagctggaca tcgtgcccct gaacagcaac 480
agcagcgagt accgcctgat caactgcaac accagcgcca tcaccaggc ctgccccaa 540
gtgagcttcg accccatccc catccactac tgcgcccccg ccggtacgc catcctgaag 600
tgcaagaaca acaccagcaa cggcaccggc ccctgccaga acgtgagcac cgtgcagtgc 660
acccacggca tcaagcccgt ggtgagcacc cccctgctgc tgaacggcag cctggccgag 720
ggcgcgaga tcatcatccg cagcaagaac ctgagcaaca acgcctacac catcatcgtg 780
cacctgaacg acagcgtgga gatcgtgtgc acccgcccc aacaacaac cgcgaaggc 840
atccgcatcg gccccggcca gaccttctac gccaccgaga acatcatcgg cgacatccgc 900
caggcccact gcaacatcag cgccggcgag tggacaagg ccgtgcagcg cgtgagcgcc 960
aagctgcgcg agcacttccc caacaagacc atcgagttcc agcccagcag cggcggcgac 1020
ctggagatca ccaccacag cttcaactgc cgccggcgag tcttctactg caacaccagc 1080
aagctgttca acagcagcta caacggcacc agctaccgcg gcaccgagag caacagcagc 1140
atcatcacc tgcctgccc catcaagcag atcatcgaca tgtggcagaa ggtgggccc 1200
gccatctacg cccccccat cgagggcaac atcacctgca gcagcagcat caccggcctg 1260
ctgctggccc gcgacggcg cctggacaac atcaccaccg agatcttccg ccccagggc 1320
ggcgacatga aggacaactg gcgcaacgag ctgtacaagt acaagggtgt ggagatcaag 1380
cccctgggcg tggcccccac cgaggccaag cgccgcgtgg tggagcgca gaagcgcgcc 1440
gtgggcatcg gcgcctgat cttcggttc ctggcgccg cggcagcaa catggcgcc 1500
gccagcatca ccctgaccgc ccaggcccg cagctgctga gcggcatcgt gcagcagcag 1560
agcaacctgc tgcgcgccat cgaggccag cagcacatgc tgcagctgac cgtgtggggc 1620
atcaagcagc tgcaggccc cgtgctggcc atcgagcgt acctgaagga ccagcagctg 1680
ctgggcatct ggggtcgag cggcaagctg atctgcacca ccaccgtgcc ctggaacagc 1740
agctggagca acaagaccca gggcgagatc tgggagaaca tgacctgat gcagtgggac 1800
aagagatca gcaactacac cggcatcatc taccgctgc tggaggagag ccagaaccag 1860
caggagcaga acgagaagga cctgctggcc ctggacagcc gcaacaacct gtggagctgg 1920
ttcaacatca gcaactggct gtggtacatc 1950

```

<210> 14

<211> 2493

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
gpl60 coding region of HIV strain AF110975

<400> 14

```

agcggcctgg gcaacctgtg ggtgaccgtg tacgacggcg tgcccgtgtg gcgcgaggcc 60
agcaccaccc tgttctgcgc cagcgacgcc aaggcctacg agaaggaggt gcacaacgtg 120
tgggccaccc acgcctgcgt gccaccgac cccaaccccc aggagatcga gctggacaac 180
gtgaccgaga acttcaacat gtggaagaac gacatggtgg accagatgca cgaggacatc 240
atcagcctgt gggaccagag cctgaagccc cgcgtaagc tgacccccct gtgctgacc 300
ctgaagtgca ccaactacag caccaactac agcaacacca tgaacgccac cagctacaac 360
aacaacacca ccgaggagat caagaactgc accttcaaca tgaccaccga gctgcgcgac 420
aagaagcagc aggtgtacgc cctgttctac aagctggaca tcgtgcccct gaacagcaac 480
agcagcgagt accgcctgat caactgcaac accagcgcca tcaccaggc ctgccccaa 540
gtgagcttcg accccatccc catccactac tgcgcccccg ccggtacgc catcctgaag 600
tgcaagaaca acaccagcaa cggcaccggc ccctgccaga acgtgagcac cgtgcagtgc 660
acccacggca tcaagcccgt ggtgagcacc cccctgctgc tgaacggcag cctggccgag 720

```



```

ggcggcgaga tcatcatccg cagcaagaac ctgagcaaca acgcctacac catcatcgtg 780
cacctgaacg acagcgtgga gatcgtgtgc acccgcccca acaacaacac ccgcaagggc 840
atccgcatcg gccccggcca gaccttctac gccaccgaga acatcatcgg cgacatccgc 900
caggcccaact gcaacatcag cgccggcgag tggacaagg ccgtgcagcg cgtgagcgcc 960
aagctgcgcg agcacttccc caacaagacc atcgagttcc agcccagcag cggcgggcgac 1020
ctggagatca ccaccacag cttcaactgc cgcgcgaggt tcttctactg caacaccagc 1080
aagctgttca acagcagcta caacggcacc agctaccgcg gcaccgagag caacagcagc 1140
atcatcaccg tgccctgccc catcaagcag atcatcgaca tgtggcagaa ggtggggccg 1200
gccatctacg ccccccccat cgagggcaac atcacctgca gcagcagcat caccggcctg 1260
ctgctggccc gcgacggcgg cctggacaac atcaccaccg agatcttccg cccccagggc 1320
ggcgacatga aggacaactg gcgcaacgag ctgtacaagt acaagggtgt ggagatcaag 1380
ccctggggcg tggccccac cgagggcaag cgccgctggg tggagcgcg gaagcgcgcc 1440
gtgggcatcg gcgcggtgat cttcggttc ctggggcgcc cggcagcaa catggggccc 1500
gccagcatca cctgaccgc ccaggcccg cagctgctga gcggcatcgt gcagcagcag 1560
agcaacctgc tgcgcccatt cgaggccag cagcacatgc tgcagctgac cgtgtggggc 1620
atcaagcagc tgcaggcccg cgtgctggcc atcgagcgt acctgaagga ccagcagctg 1680
ctgggcatct ggggctgcag cggcaagctg atctgcacca ccaccgtgcc ctggaacagc 1740
agctggagca acaagaccca gggcgagatc tgggagaaca tgacctggat gcagtgggac 1800
aaggagatca gcaactacac cggcatcatc taccgctgc tggaggagag ccagaaccag 1860
caggagcaga acgagaagga cctgctggcc ctggacagcc gcaacaacct gtggagctgg 1920
ttcaacatca gcaactggct gtggtacatc aagatcttca tcatgatcgt gggcggcctg 1980
atcggcctgc gcatcatctt cgccgtgctg agcatcgtga accgctgctg ccagggttac 2040
agccccctga gcttcagac cctgaccccc aacccccgcg gcctggaccg cctggggccg 2100
atcgaggagg agggcgcgga gcaggaccgc gaccgcagca tccgcctggt gcagggttc 2160
ctggccctgg cctgggacga cctgcgcagc ctgtgcctgt tcagctacca ccgcctgctg 2220
gacctgatcc tggtgaccgc ccgctgggtg gagctgctgg gccgcagcag ccccccgccg 2280
ctgcagcgcg gctgggaggc cctgaagtac ctgggcagcc tggtgagta ctggggcctg 2340
gagctgaaga agagcgccac cagcctgctg gacagcatcg ccacgcccgt ggccgagggc 2400
accgaccgca tcatcgaggt gatccagcgc atctaccgcg cttctgcaa catccccgc 2460
cgctgcgccc agggcttcga ggccgcccgt cag
2493

```

<210> 15

<211> 2565

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
signal sequence and gp160 coding region of HIV
strain AF110975

<400> 15

```

atgcgcgtgc gcggcatcct gcgcagctgg cagcagtggt ggatctgggg catcctgggc 60
ttctggatct gcagcggcct gggcaacctg tgggtgaccg tgtacgacgg cgtgcccgtg 120
tggcgcgagg ccagcaccac cctgttctgc gccagcgacg ccaaggccta cgagaaggag 180
gtgcacaacg tgtggggccac ccacgcctgc gtgcccaccg accccaacct ccaggagatc 240
gagctggaca acgtgaccga gaacttcaac atgtggaaga acgacatggt ggaccagatg 300
caccagagaca tcatcagcct gtgggaccag agcctgaagc cccgcgtgaa gctgaccccc 360
ctgtgcgtga ccctgaagtg caccaactac agcaccaact acagcaacac catgaacgcc 420
accagctaca acaacaacac caccgaggag atcaagaact gcaccttcaa catgaccacc 480
gagctgcgcg acaagaagca gcaggtgtac gccctgttct acaagctgga catcgtgccc 540
ctgaacagca acagcagcga gtaccgcctg atcaactgca acaccagcgc catcaccagc 600
gcctgccccg aggtgagctt cgaccccatc cccatccact actgcgcccc cgccggctac 660
gccatcctga agtgcaagaa caacaccagc aacggcaccg gccctgcca gaacgtgagc 720
accgtgcagt gcaccacagg catcaagccc gtggtgagca cccccctgct gctgaacggc 780
agcctggccc agggcgcgga gatcatcatc cgcagcaaga acctgagcaa caacgcctac 840

```

accatcatcg	tgcacctgaa	cgacagcggtg	gagatcggtg	gcacccgccc	caacaacaac	900
acccgcaagg	gcatccgcat	cgggcccggc	cagaccttct	acgccaccga	gaacatcatc	960
ggcgacatcc	gccaggccca	ctgcaacatc	agcgccggcg	agtggaacaa	ggccgtgcag	1020
cgcgtagagcg	ccaagctgcg	cgagcacttc	cccaacaaga	ccatcgagtt	ccagcccagc	1080
agcggcggcg	acctggagat	caccacccac	agcttcaact	gccgcggcga	gttctttctac	1140
tgcaacacca	gcaagctgtt	caacagcagc	tacaacggca	ccagctaccg	cggcaccgag	1200
agcaacagca	gcatcatcac	cctgccctgc	cgcacaaagc	agatcatcga	catgtggcag	1260
aaggtggggc	gcgccatcta	cgcccccccc	atcgagggca	acatcacctg	cagcagcagc	1320
atcaccggcc	tgctgctggc	ccgcgacggc	ggcctggaca	acatcaccac	cgagatcttc	1380
cgcccccagg	gcggcgacat	gaaggacaac	tggcgcaacg	agctgtacaa	gtacaaggtg	1440
gtggagatca	agccccctggg	cgtggccccc	accgaggcca	agcgccgcgt	ggtggagcgc	1500
gagaagcgcg	ccgtgggcct	cggcgcggtg	atcttcggct	tcctggggcg	cgccggcagc	1560
aacatggggc	ccgccagcat	cacctgacc	gcccaggccc	gccagctgct	gagcggcatc	1620
gtgcagcagc	agagcaacct	gctgcgcgcc	atcgaggccc	agcagcacat	gctgcagctg	1680
accgtgtggg	gcatcaagca	gctgcaggcc	cgctgtctgg	ccatcgagcg	ctacctgaag	1740
gaccagcagc	tgctgggcat	ctggggctgc	agcggcaagc	tgatctgcac	caccaccgtg	1800
ccctggaaca	gcagctggag	caacaagacc	cagggcgaga	tctgggagaa	catgacctgg	1860
atgcagtggg	acaaggagat	cagcaactac	accggcatca	tctaccgcct	gctggaggag	1920
agccagaacc	agcaggagca	gaacgagaag	gacctgctgg	ccctggacag	ccgcaacaac	1980
ctgtggagct	ggttcaacat	cagcaactgg	ctgtggtaca	tcaagatctt	catcatgatc	2040
gtgggcggcc	tgatcggcct	gcgcacatc	ttcgccgtgc	tgagcatcgt	gaaccgcgtg	2100
cgccagggct	acagccccct	gagcttccag	accctgacct	ccaacccccg	cggcctggac	2160
cgccctgggc	gcatcgagga	ggagggcggc	gagcaggacc	gcgaccgcag	catccgcctg	2220
gtgcagggct	tcctggccct	ggcctgggac	gacctgcgca	gcctgtgcct	gttcagctac	2280
caccgcctgc	gcgacctgat	cctggtgacc	gcccgcgtgg	tggagctgct	gggcccgcagc	2340
agcccccgcg	gcctgcagcg	cggctgggag	gccctgaagt	acctgggcag	cctggtgcag	2400
tactggggcc	tggagctgaa	gaagagcgcc	accagcctgc	tggacagcat	cgccatcgcc	2460
gtggccgagg	gcaccgaccg	catcatcgag	gtgatccagc	gcactaccg	cgccttctgc	2520
aacatcccc	gccgcgtgcg	ccagggcttc	gaggccgccc	tgag		2565

<210> 16

<211> 1056

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic a
gp41 coding region of HIV strain AF110975

<400> 16

gccgtgggca	tcggcgccgt	gatcttcggc	ttcctgggcg	ccgcggcgag	caacatgggc	60
gccgccagca	tcaccctgac	cgcccaggcc	cgccagctgc	tgagcggcat	cgtgcagcag	120
cagagcaacc	tgctgcgcgc	catcgaggcc	cagcagcaca	tgctgcagct	gaccgtgtgg	180
ggcatcaagc	agctgcaggc	ccgcgtgctg	gccatcgagc	gtacctgaa	ggaccagcag	240
atgctgggca	tctggggctg	cagcggcaag	ctgatctgca	ccaccaccgt	gccctggaac	300
agcagctgga	gcaacaagac	ccagggcgag	atctgggaga	acatgacctg	gatgcagtgg	360
gacaaggaga	tcagcaacta	caccggcatc	atctaccgcc	tgctggagga	gagccagaac	420
cagcaggagc	agaacgagaa	ggacctgctg	gccctggaca	gccgcaacaa	cctgtggagc	480
tggttcaaca	tcagcaactg	gctgtggtac	atcaagatct	tcacatcatg	cgtgggcggc	540
ctgatcggcc	tgcgcatcat	cttcgccgtg	ctgagcatcg	tgaaccgcgt	gcgccagggc	600
tacagcccc	tgagcttcca	gacctgacc	cccaaccccc	gcggcctgga	ccgcctgggc	660
cgcacagagg	aggagggcg	cgagcaggac	cgcgaccgca	gcacccgcct	ggtgcagggc	720
ttcctggccc	tggcctggga	cgacctgcgc	agcctgtgcc	tgttcagcta	ccaccgcctg	780
cgcgacctga	tcctggtgac	cgcccgcgtg	gtggagctgc	tgggcccgcag	cagcccccg	840
ggcctgcagc	gcggctggga	ggcctgaag	tacctgggca	gcctggtgca	gtactggggc	900
ctggagctga	agaagagcgc	caccagcctg	ctggacagca	tcgccatcgc	cgtggccgag	960

ggcaccgacc gcatcatcga ggtgatccag cgcacttacc gcgccttctg caacatcccc 1020
cgccgcgtgc gccagggctt cgaggccgcc ctgcag 1056

<210> 17
<211> 492
<212> PRT
<213> Human immunodeficiency virus

<400> 17
Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Ala Trp
1 5 10 15
Glu Arg Ile Arg Leu Arg Pro Gly Gly Lys Lys Cys Tyr Met Met Lys
20 25 30
His Leu Val Trp Ala Ser Arg Glu Leu Glu Lys Phe Ala Leu Asn Pro
35 40 45
Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Arg Gln Leu
50 55 60
His Pro Ala Leu Gln Thr Gly Ser Glu Glu Leu Lys Ser Leu Phe Asn
65 70 75 80
Thr Val Ala Thr Leu Tyr Cys Val His Glu Lys Ile Glu Val Arg Asp
85 90 95
Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Cys Gln
100 105 110
Gln Lys Ile Gln Gln Ala Glu Ala Ala Asp Lys Gly Lys Val Ser Gln
115 120 125
Asn Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Ala
130 135 140
Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys
145 150 155 160
Ala Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly
165 170 175
Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His
180 185 190
Gln Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala
195 200 205
Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala Pro Gly
210 215 220
Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Thr
225 230 235 240
Leu Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile Pro Val
245 250 255

Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val
 260 265 270
 Arg Met Tyr Ser Pro Val Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys
 275 280 285
 Glu Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala
 290 295 300
 Glu Gln Ser Thr Gln Glu Val Lys Asn Trp Met Thr Asp Thr Leu Leu
 305 310 315 320
 Val Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly
 325 330 335
 Pro Gly Ala Ser Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly
 340 345 350
 Gly Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Ala
 355 360 365
 Asn Thr Ser Val Met Met Gln Lys Ser Asn Phe Lys Gly Pro Arg Arg
 370 375 380
 Ile Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn
 385 390 395 400
 Cys Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly
 405 410 415
 His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys
 420 425 430
 Ile Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg
 435 440 445
 Pro Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr
 450 455 460
 Thr Pro Gly Gln Lys Gln Glu Ser Lys Asp Arg Glu Thr Leu Thr Ser
 465 470 475 480
 Leu Lys Ser Leu Phe Gly Asn Asp Pro Leu Ser Gln
 485 490

<210> 18

<211> 81

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic
 signal sequence of HIV strain AF110968

<400> 18

atgcgcgtga tgggcatcct gaagaactac cagcagtggg ggatgtgggg catcctgggc 60
 ttctggatgc tgatcatcag c 81

<210> 19
 <211> 72
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: synthetic
 signal sequence of HIV strain AF110975

<400> 19
 atgcgcgtgc gcggcacatcct gcgcagctgg cagcagtggg ggatctgggg catcctgggc 60
 ttctggatct gc 72

<210> 20
 <211> 1479
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: synthetic Gag
 coding sequence of HIV strain AF110965

<400> 20
 atgggcgccc ggcgcagcat cctgcgcggc ggcaagctgg acgcctggga gcgcacccgc 60
 ctgcgccccg gcggaagaa gtgctacatg atgaagcacc tgggtgtggc cagccgcgag 120
 ctggagaagt tcgccctgaa ccccgccctg ctggagacca gcgagggtg caagcagatc 180
 atccgcccgc tgcaccccgc cctgcagacc ggcagcgagg agctgaagag cctgttcaac 240
 accgtggcca cccgtgactg cgtgcacgag aagatcgagg tgcgcgacac caaggaggcc 300
 ctggacaaga tcgaggagga gcagaacaag tgccagcaga agatccagca ggccgaggcc 360
 gccgacaagg gcaaggtgag ccagaactac cccatcgtgc agaacctgca gggccagatg 420
 gtgcaccagg ccatcagccc ccgcaccctg aacgcctggg tgaaggtgat cgaggagaag 480
 gccttcagcc ccgagggtgat ccccatgttc accgcctga gcgaggggcg cccccccag 540
 gacctgaaca ccatgctgaa caccgtgggc ggccaccagg ccgccatgca gatgctgaag 600
 gacaccatca acgaggaggc cgccgagtgg gaccgcgtgc acccctgca cgccggcccc 660
 atcgcccccg gccagatgcg cgagccccgc ggcagcgaca tcgccggcac caccagcacc 720
 ctgcaggagc agatgcctg gatgaccagc aacccccca tccccgtggg cgacatctac 780
 aagcgctgga tcatcctggg cctgaacaag atcgtgcgca tgtacagccc cgtgagcatc 840
 ctggacatca agcaggggccc caaggagccc ttccgcgact acgtggaccg cttcttcaag 900
 accctgcgcg ccgagcagag caccagagg gtgaagaact ggatgaccga caccctgctg 960
 gtgcagaacg ccaaccccga ctgcaagacc atcctgcgcg ccctggggcc cgccgcccagc 1020
 ctggaggaga tgatgaccgc ctgccagggc gtggggcgcc ccagccacaa ggcccgctg 1080
 ctggccgagg ccatgagcca ggccaacacc agcgtgatga tgcagaagag caacttcaag 1140
 ggccccgcc gcacgtgaa gtgcttcaac tgcggcaagg agggccacat cgcccgcaac 1200
 tgccgcgccc ccgcaagaa gggctgctgg aagtgcggca aggaggcca ccagatgaag 1260
 gactgcaccg agcgccaggc caacttcctg ggcaagatct ggccagcca caagggccgc 1320
 cccggcaact tcctgcagag ccgccccgag cccaccgccc ccccgccga gagcttccgc 1380
 ttcgaggaga ccacccccg ccagaagcag gagagcaagg accgcgagac cctgaccagc 1440
 ctgaagagcc tgttcgcaa cgacccctg agccagtaa 1479

<210> 21
 <211> 1509
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: synthetic Gag
coding sequence of HIV strain AF110967

<400> 21

```
atgggcgccc gcgccagcat cctgcgcggc gagaagctgg acaagtggga gaagatccgc 60
ctgcgccccg gcggaagaa gcactacatg ctgaagcacc tgggtgtggg cagccgcgag 120
ctggagggct tcgccctgaa ccccggcctg ctggagaccg ccgagggctg caagcagatc 180
atgaagcagc tgcagccgc cctgcagacc ggcaccgagg agctgcgcag cctgtacaac 240
accgtggcca ccctgtactg cgtgcacgcc ggcatcgagg tgcgcgacac caaggaggcc 300
ctggacaaga tcgaggagga gcagaacaag agccagcaga agaccagca ggccaaggag 360
gccagggcca aggtgagcca gaactacccc atcgtgcaga acctgcaggg ccagatggtg 420
caccaggcca tcagcccccg caccctgaac gcctgggtga aggtgatcga ggagaaggcc 480
ttcagccccg aggtgatccc catgttcacc gccctgagcg agggcgccac ccccaggac 540
ctgaacacca tgctgaacac cgtgggcggc caccaggccg ccatgcagat gctgaaggac 600
accatcaacg aggaggccgc cgagtgggac cgctgcacc ccgtgcaggc cggccccgtg 660
gcccccgccc agatgcgcga cccccgcggc agcgacatcg ccggcgccac cagcaccctg 720
caggagcaga tcgcctggat gaccagcaac cccccgtgc ccgtgggcga catctacaag 780
cgctggatca tcctgggcct gaacaagatc gtgcgcatgt acagccccgt gagcatcctg 840
gacatccgcc agggcccaa ggagcccttc cgcgactacg tggaccgctt cttcaagacc 900
ctgcgcgccg agcaggccac ccaggacgtg aagaactgga tgaccgagac cctgctggtg 960
cagaacgcca accccgactg caagaccatc ctgcgcgccc tgggccccgg cgccaccctg 1020
gaggagatga tgaccgcctg ccagggcgtg ggcgcccccg gccacaaggc ccgcgtgctg 1080
gccgaggcca tgagccaggc caacagcgtg aacatcatga tgcagaagag caacttcaag 1140
ggcccccgcc gcaacgtgaa gtgcttcaac tgcggcaagg agggccacat cgccaagaac 1200
tgccgcgccc cccgcaagaa gggctgctgg aagtgcggca aggagggcca ccagatgaag 1260
gactgcaccg agcgccaggc caacttctct ggcaagatct ggcccagcca caagggccgc 1320
cccggcaact tcctgcagaa ccgcagcgag cccgcgccc ccaccgtgcc caccgcccc 1380
cccgccgaga gttccgctt cgaggagacc accccgccc ccaagcagga gcccaaggac 1440
cgcgagccct accgcgagcc cctgaccgcc ctgcgcgacc tgttcggcag cggccccctg 1500
agccagtaa
```

<210> 22

<211> 502

<212> PRT

<213> Human immunodeficiency virus

<400> 22

```
Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Glu Lys Leu Asp Lys Trp
  1             5             10             15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys
      20             25             30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Gly Phe Ala Leu Asn Pro
      35             40             45

Gly Leu Leu Glu Thr Ala Glu Gly Cys Lys Gln Ile Met Lys Gln Leu
      50             55             60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Tyr Asn
      65             70             75             80

Thr Val Ala Thr Leu Tyr Cys Val His Ala Gly Ile Glu Val Arg Asp
      85             90             95
```

Thr	Lys	Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Gln	
			100					105					110			
Gln	Lys	Thr	Gln	Gln	Ala	Lys	Glu	Ala	Asp	Gly	Lys	Val	Ser	Gln	Asn	
		115					120					125				
Tyr	Pro	Ile	Val	Gln	Asn	Leu	Gln	Gly	Gln	Met	Val	His	Gln	Ala	Ile	
	130					135					140					
Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Ile	Glu	Glu	Lys	Ala	
145					150					155					160	
Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Thr	Ala	Leu	Ser	Glu	Gly	Ala	
			165						170					175		
Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His	Gln	
			180					185					190			
Ala	Ala	Met	Gln	Met	Leu	Lys	Asp	Thr	Ile	Asn	Glu	Glu	Ala	Ala	Glu	
		195					200					205				
Trp	Asp	Arg	Leu	His	Pro	Val	Gln	Ala	Gly	Pro	Val	Ala	Pro	Gly	Gln	
	210					215					220					
Met	Arg	Asp	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Ala	Thr	Ser	Thr	Leu	
225					230					235					240	
Gln	Glu	Gln	Ile	Ala	Trp	Met	Thr	Ser	Asn	Pro	Pro	Val	Pro	Val	Gly	
			245						250					255		
Asp	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val	Arg	
			260					265					270			
Met	Tyr	Ser	Pro	Val	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly	Pro	Lys	Glu	
		275					280					285				
Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Phe	Lys	Thr	Leu	Arg	Ala	Glu	
	290					295					300					
Gln	Ala	Thr	Gln	Asp	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr	Leu	Leu	Val	
305					310					315					320	
Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Arg	Ala	Leu	Gly	Pro	
				325					330					335		
Gly	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly	Val	Gly	Gly	
			340					345					350			
Pro	Gly	His	Lys	Ala	Arg	Val	Leu	Ala	Glu	Ala	Met	Ser	Gln	Ala	Asn	
		355					360					365				
Ser	Val	Asn	Ile	Met	Met	Gln	Lys	Ser	Asn	Phe	Lys	Gly	Pro	Arg	Arg	
	370					375					380					
Asn	Val	Lys	Cys	Phe	Asn	Cys	Gly	Lys	Glu	Gly	His	Ile	Ala	Lys	Asn	
385					390					395					400	

145		150		155		160
Glu Leu Arg Asp	Arg Lys Gln Glu Val	His Ala Leu Phe Tyr Arg Leu				
	165	170			175	
Asp Val Val Pro	Leu Gln Gly Asn Asn Ser Asn Glu Tyr Arg Leu Ile					
	180	185			190	
Asn Cys Asn Thr	Ser Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe					
	195	200			205	
Asp Pro Ile Pro	Ile His Tyr Cys Thr Pro Ala Gly Tyr Ala Ile Leu					
	210	215			220	
Lys Cys Asn Asn	Gln Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val					
	225	230			235	240
Ser Ser Val Gln	Cys Ala His Gly Ile Lys Pro Val Val Ser Thr Gln					
	245	250			255	
Leu Leu Leu Asn	Gly Ser Leu Ala Lys Gly Glu Ile Ile Ile Arg Ser					
	260	265			270	
Glu Asn Leu Ala	Asn Asn Ala Lys Ile Ile Ile Val Gln Leu Asn Lys					
	275	280			285	
Pro Val Lys Ile	Val Cys Val Arg Pro Asn Asn Asn Thr Arg Lys Ser					
	290	295			300	
Val Arg Ile Gly	Pro Gly Gln Thr Phe Tyr Ala Thr Gly Glu Ile Ile					
	305	310			315	320
Gly Asp Ile Arg	Gln Ala Tyr Cys Ile Ile Asn Lys Thr Glu Trp Asn					
	325	330			335	
Ser Thr Leu Gln	Gly Val Ser Lys Lys Leu Glu Glu His Phe Ser Lys					
	340	345			350	
Lys Ala Ile Lys	Phe Glu Pro Ser Ser Gly Gly Asp Leu Glu Ile Thr					
	355	360			365	
Thr His Ser Phe	Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asp Thr Ser					
	370	375			380	
Gln Leu Phe Asn	Ser Thr Tyr Ser Pro Ser Phe Asn Gly Thr Glu Asn					
	385	390			395	400
Lys Leu Asn Gly	Thr Ile Thr Ile Thr Cys Arg Ile Lys Gln Ile Ile					
	405	410			415	
Asn Met Trp Gln	Lys Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala					
	420	425			430	
Gly Asn Leu Thr	Cys Glu Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg					
	435	440			445	
Asp Gly Gly Lys	Thr Gly Pro Asn Asp Thr Glu Ile Phe Arg Pro Gly					

450					455					460					
Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	Asn	Glu	Leu	Tyr	Lys	Tyr	Lys
465					470					475					480
Val	Val	Glu	Ile	Lys	Pro	Leu	Gly	Val	Ala	Pro	Thr	Glu	Ala	Lys	Arg
				485					490					495	
Arg	Val	Val	Glu	Arg	Glu	Lys	Arg	Ala	Val	Gly	Ile	Gly	Ala	Val	Phe
			500					505					510		
Leu	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Ile
		515					520					525			
Thr	Leu	Thr	Val	Gln	Ala	Arg	Leu	Leu	Leu	Ser	Gly	Ile	Val	Gln	Gln
	530					535					540				
Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	Gln
545					550					555					560
Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Thr	Arg	Ile	Leu	Ala	Val
				565					570					575	
Glu	Arg	Tyr	Leu	Lys	Asp	Gln	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	Ser
			580					585					590		
Gly	Lys	Leu	Ile	Cys	Thr	Thr	Ala	Val	Pro	Trp	Asn	Ser	Ser	Trp	Ser
		595					600					605			
Asn	Arg	Ser	His	Asp	Glu	Ile	Trp	Asp	Asn	Met	Thr	Trp	Met	Gln	Trp
						615					620				
Asp	Arg	Glu	Ile	Asn	Asn	Tyr	Thr	Asp	Thr	Ile	Tyr	Arg	Leu	Leu	Glu
625					630					635					640
Glu	Ser	Gln	Asn	Gln	Gln	Glu	Lys	Asn	Glu	Lys	Asp	Leu	Leu	Ala	Leu
				645					650					655	
Asp	Ser	Trp	Gln	Asn	Leu	Trp	Asn	Trp	Phe	Ser	Ile	Thr	Asn	Trp	Leu
			660					665					670		
Trp	Tyr	Ile	Lys	Ile	Phe	Ile	Met	Ile	Val	Gly	Gly	Leu	Ile	Gly	Leu
		675					680					685			
Arg	Ile	Ile	Phe	Ala	Val	Leu	Ser	Ile	Val	Asn	Arg	Val	Arg	Gln	Gly
	690					695					700				
Tyr	Ser	Pro	Leu	Pro	Phe	Gln	Thr	Leu	Thr	Pro	Asn	Pro	Arg	Glu	Pro
705					710					715					720
Asp	Arg	Leu	Gly	Arg	Ile	Glu	Glu	Glu	Gly	Gly	Glu	Gln	Asp	Arg	Gly
				725					730					735	
Arg	Ser	Ile	Arg	Leu	Val	Ser	Gly	Phe	Leu	Ala	Leu	Ala	Trp	Asp	Asp
			740					745					750		

Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile
 755 760 765
 Leu Ile Ala Ala Arg Val Leu Glu Leu Leu Gly Gln Arg Gly Trp Glu
 770 775 780
 Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln Tyr Trp Gly Leu Glu Leu
 785 790 795 800
 Lys Lys Ser Ala Ile Ser Leu Leu Asp Thr Ile Ala Ile Ala Val Ala
 805 810 815
 Glu Gly Thr Asp Arg Ile Ile Glu Phe Ile Gln Arg Ile Cys Arg Ala
 820 825 830
 Ile Arg Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala Ala Leu
 835 840 845

Gln

<210> 24
 <211> 855
 <212> PRT
 <213> Human immunodeficiency virus

<400> 24
 Met Arg Val Arg Gly Ile Leu Arg Ser Trp Gln Gln Trp Trp Ile Trp
 1 5 10 15
 Gly Ile Leu Gly Phe Trp Ile Cys Ser Gly Leu Gly Asn Leu Trp Val
 20 25 30
 Thr Val Tyr Asp Gly Val Pro Val Trp Arg Glu Ala Ser Thr Thr Leu
 35 40 45
 Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val His Asn Val
 50 55 60
 Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro Gln Glu Ile
 65 70 75 80
 Glu Leu Asp Asn Val Thr Glu Asn Phe Asn Met Trp Lys Asn Asp Met
 85 90 95
 Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp Gln Ser Leu
 100 105 110
 Lys Pro Arg Val Lys Leu Thr Pro Leu Cys Val Thr Leu Lys Cys Thr
 115 120 125
 Asn Tyr Ser Thr Asn Tyr Ser Asn Thr Met Asn Ala Thr Ser Tyr Asn
 130 135 140
 Asn Asn Thr Thr Glu Glu Ile Lys Asn Cys Thr Phe Asn Met Thr Thr

145					150						155					160
Glu	Leu	Arg	Asp	Lys	Lys	Gln	Gln	Val	Tyr	Ala	Leu	Phe	Tyr	Lys	Leu	
				165					170					175		
Asp	Ile	Val	Pro	Leu	Asn	Ser	Asn	Ser	Ser	Glu	Tyr	Arg	Leu	Ile	Asn	
			180					185					190			
Cys	Asn	Thr	Ser	Ala	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	Asp	
		195					200					205				
Pro	Ile	Pro	Ile	His	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr	Ala	Ile	Leu	Lys	
	210					215					220					
Cys	Lys	Asn	Asn	Thr	Ser	Asn	Gly	Thr	Gly	Pro	Cys	Gln	Asn	Val	Ser	
225					230					235						240
Thr	Val	Gln	Cys	Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser	Thr	Pro	Leu	
				245					250					255		
Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Gly	Gly	Glu	Ile	Ile	Ile	Arg	Ser	
			260					265						270		
Lys	Asn	Leu	Ser	Asn	Asn	Ala	Tyr	Thr	Ile	Ile	Val	His	Leu	Asn	Asp	
		275					280					285				
Ser	Val	Glu	Ile	Val	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Gly	
		290				295					300					
Ile	Arg	Ile	Gly	Pro	Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Glu	Asn	Ile	Ile	
305					310					315					320	
Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn	Ile	Ser	Ala	Gly	Glu	Trp	Asn	
				325					330					335		
Lys	Ala	Val	Gln	Arg	Val	Ser	Ala	Lys	Leu	Arg	Glu	His	Phe	Pro	Asn	
			340					345					350			
Lys	Thr	Ile	Glu	Phe	Gln	Pro	Ser	Ser	Gly	Gly	Asp	Leu	Glu	Ile	Thr	
		355					360					365				
Thr	His	Ser	Phe	Asn	Cys	Arg	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	
	370					375					380					
Lys	Leu	Phe	Asn	Ser	Ser	Tyr	Asn	Gly	Thr	Ser	Tyr	Arg	Gly	Thr	Glu	
385					390					395					400	
Ser	Asn	Ser	Ser	Ile	Ile	Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	
				405					410					415		
Asp	Met	Trp	Gln	Lys	Val	Gly	Arg	Ala	Ile	Tyr	Ala	Pro	Pro	Ile	Glu	
			420					425					430			
Gly	Asn	Ile	Thr	Cys	Ser	Ser	Ser	Ile	Thr	Gly	Leu	Leu	Leu	Ala	Arg	
		435					440					445				
Asp	Gly	Gly	Leu	Asp	Asn	Ile	Thr	Thr	Glu	Ile	Phe	Arg	Pro	Gln	Gly	

450					455					460					
Gly	Asp	Met	Lys	Asp	Asn	Trp	Arg	Asn	Glu	Leu	Tyr	Lys	Tyr	Lys	Val
465					470					475					480
Val	Glu	Ile	Lys	Pro	Leu	Gly	Val	Ala	Pro	Thr	Glu	Ala	Lys	Arg	Arg
				485					490					495	
Val	Val	Glu	Arg	Glu	Lys	Arg	Ala	Val	Gly	Ile	Gly	Ala	Val	Ile	Phe
			500					505					510		
Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Asn	Met	Gly	Ala	Ala	Ser	Ile	Thr
		515					520					525			
Leu	Thr	Ala	Gln	Ala	Arg	Gln	Leu	Leu	Ser	Gly	Ile	Val	Gln	Gln	Gln
	530					535					540				
Ser	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Met	Leu	Gln	Leu
545					550					555					560
Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Ala	Arg	Val	Leu	Ala	Ile	Glu
				565					570					575	
Arg	Tyr	Leu	Lys	Asp	Gln	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	Ser	Gly
			580					585					590		
Lys	Leu	Ile	Cys	Thr	Thr	Thr	Val	Pro	Trp	Asn	Ser	Ser	Trp	Ser	Asn
		595					600					605			
Lys	Thr	Gln	Gly	Glu	Ile	Trp	Glu	Asn	Met	Thr	Trp	Met	Gln	Trp	Asp
	610					615					620				
Lys	Glu	Ile	Ser	Asn	Tyr	Thr	Gly	Ile	Ile	Tyr	Arg	Leu	Leu	Glu	Glu
625					630					635					640
Ser	Gln	Asn	Gln	Gln	Glu	Gln	Asn	Glu	Lys	Asp	Leu	Leu	Ala	Leu	Asp
				645					650					655	
Ser	Arg	Asn	Asn	Leu	Trp	Ser	Trp	Phe	Asn	Ile	Ser	Asn	Trp	Leu	Trp
			660					665					670		
Tyr	Ile	Lys	Ile	Phe	Ile	Met	Ile	Val	Gly	Gly	Leu	Ile	Gly	Leu	Arg
		675					680					685			
Ile	Ile	Phe	Ala	Val	Leu	Ser	Ile	Val	Asn	Arg	Val	Arg	Gln	Gly	Tyr
	690					695					700				
Ser	Pro	Leu	Ser	Phe	Gln	Thr	Leu	Thr	Pro	Asn	Pro	Arg	Gly	Leu	Asp
705					710					715					720
Arg	Leu	Gly	Arg	Ile	Glu	Glu	Glu	Gly	Gly	Glu	Gln	Asp	Arg	Asp	Arg
				725					730					735	
Ser	Ile	Arg	Leu	Val	Gln	Gly	Phe	Leu	Ala	Leu	Ala	Trp	Asp	Asp	Leu
			740					745					750		

Arg Ser Leu Cys Leu Phe Ser Tyr His Arg Leu Arg Asp Leu Ile Leu
755 760 765

Val Thr Ala Arg Val Val Glu Leu Leu Gly Arg Ser Ser Pro Arg Gly
770 775 780

Leu Gln Arg Gly Trp Glu Ala Leu Lys Tyr Leu Gly Ser Leu Val Gln
785 790 795 800

Tyr Trp Gly Leu Glu Leu Lys Lys Ser Ala Thr Ser Leu Leu Asp Ser
805 810 815

Ile Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Val Ile
820 825 830

Gln Arg Ile Tyr Arg Ala Phe Cys Asn Ile Pro Arg Arg Val Arg Gln
835 840 845

Gly Phe Glu Ala Ala Leu Gln
850 855

<210> 25
<211> 20
<212> PRT
<213> Human immunodeficiency virus

<400> 25
Asp Ile Lys Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg
1 5 10 15

Phe Phe Lys Thr
20

<210> 26
<211> 60
<212> DNA
<213> Human immunodeficiency virus

<400> 26
gacataaaac aaggaccaa agagcccttt agagactatg tagaccggtt ctttaaaacc 60

<210> 27
<211> 20
<212> PRT
<213> Human immunodeficiency virus

<400> 27
Asp Ile Arg Gln Gly Pro Lys Glu Pro Phe Arg Asp Tyr Val Asp Arg
1 5 10 15

Phe Phe Lys Thr
20

<210> 28
 <211> 47
 <212> PRT
 <213> Human immunodeficiency virus

<400> 28
 Thr Ile Thr Ile Thr Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln
 1 5 10 15
 Lys Val Gly Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Leu Thr
 20 25 30
 Cys Glu Ser Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly
 35 40 45

<210> 29
 <211> 48
 <212> PRT
 <213> Human immunodeficiency virus

<400> 29
 Ser Ile Ile Thr Leu Pro Cys Arg Ile Lys Gln Ile Ile Asp Met Trp
 1 5 10 15
 Gln Lys Val Gly Arg Ala Ile Tyr Ala Pro Pro Ile Glu Gly Asn Ile
 20 25 30
 Thr Cys Ser Ser Ser Ile Thr Gly Leu Leu Leu Ala Arg Asp Gly Gly
 35 40 45

<210> 30
 <211> 2469
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: PR975(+)

<400> 30
 gtcgacgccca ccatggccga ggccatgagc caggccacca gcgccaacat cctgatgcag 60
 cgcagcaact tcaagggccc caagcgcata atcaagtgtc tcaactgcgg caaggagggc 120
 cacatcgccc gcaactgccg cgccccccgc aagaagggct gctggaagtg cggcaaggag 180
 ggccaccaga tgaaggactg caccgagcgc caggccaact tcttccgcga ggacctggcc 240
 ttccccccag gcaagggccc cgagttcccc agcagacaga accgcgcca cagccccacc 300
 agccgcgagc tgcaggtgcg cggcgacaac ccccgagcgc aggccggcgc cgagcgccag 360
 ggcaaccctga acttccccca gatcaccctg tggcagcgcc ccctggtgag catcaagggtg 420
 ggcgggccaga tcaaggaggc cctgctggac accggcgccg acgacaccgt gctggaggag 480
 atgagcctgc ccggcaagtg gaagcccaag atgatcggcg gcatcggcgg cttcatcaag 540
 gtgcgccagt acgaccagat cctgatcgag atctgcggca agaaggccat cggcaccgtg 600
 ctgatcgccc ccacccccgt gaacatcatc ggccgcaaca tgctgacca gctgggctgc 660
 accctgaact tccccatcag ccccatcgag accgtgcccg tgaagctgaa gcccggcatg 720
 gacggcccca aggtgaagca gtggcccctg accgaggaga agatcaaggc cctgaccgcc 780
 atctgcgagg agatggagaa ggagggcaag atcaccaaga tcggccccga gaacccttac 840
 aacacccccg tgttcgccat caagaagaag gacagacca agtggcgcaa gctggtggac 900
 ttccgcgagc tgaacaagcg caccaggac ttctgggagg tgcagctggg catccccac 960

cccgccggcc	tgaagaagaa	gaagagcgtg	accgtgctgg	acgtgggcca	cgcctacttc	1020
agcgtgcccc	tggacgagga	cttccgcaag	tacaccgcct	tcaccatccc	cagcatcaac	1080
aacgagaccc	ccggcatccg	ctaccagtac	aacgtgctgc	cccagggctg	gaagggcagc	1140
cccagcatct	tccagagcag	catgaccaag	atcctggagc	ccttccgcgc	ccgcaacccc	1200
gagatcgtga	tctaccagta	catggacgac	ctgtacgtgg	gcagcgacct	ggagatcggc	1260
cagcaccgcg	ccaagatcga	ggagctgcgc	aagcacctgc	tgcgctgggg	cttcaccacc	1320
cccgacaaga	agcaccagaa	ggagccccc	ttcctgtgga	tgggctacga	gctgcacccc	1380
gacaagtgga	ccgtgcagcc	catcgagctg	cccagagaag	agagctggac	cgtgaacgac	1440
atccagaagc	tgggtgggcaa	gctgaactgg	gccagccaga	tctaccccgg	catcaaggtg	1500
cgccagctgt	gcaagctgct	gcgcggcgcc	aaggccctga	ccgacatcgt	gcccctgacc	1560
gaggaggccg	agctggagct	ggccgagaac	cgcgagatcc	tgcgcgagcc	cgtgcacggc	1620
gtgtactacg	accccagcaa	ggacctggtg	gccgagatcc	agaagcaggg	ccacgaccag	1680
tggacctacc	agatctacca	ggagcccttc	aagaacctga	agaccggcaa	gtacgccaa	1740
atgcgcaccg	cccacaccaa	cgacgtgaag	cagctgaccg	aggccgtgca	gaagatcgcc	1800
atggagagca	tcgtgatctg	gggcaagacc	cccaagttcc	gcctgcccac	ccagaaggag	1860
acctgggaga	cctggtggac	cgactactgg	caggccacct	ggatccccga	gtgggagttc	1920
gtgaacaccc	cccccttgt	gaagctgtgg	taccagctgg	agaaggagcc	catcatcgcc	1980
gccgagacct	tctacgtgga	cgcgccgcgc	aaccgcgaga	ccaagatcgg	caaggccggc	2040
tacgtgaccg	accggggccg	gcagaagatc	gtgagcctga	ccgagaccac	caaccagaag	2100
accgagctgc	aggccatcca	gctggccctg	caggacagcg	gcagcgaggt	gaacatcgtg	2160
accgacagcc	agtagccct	gggcatcatc	caggcccagc	ccgacaagag	cgagagcgag	2220
ctggtgaacc	agatcatcga	gcagctgatc	aagaaggaga	aggtgtacct	gagctgggtg	2280
cccgcccaca	agggcatcgg	cggcaacgag	cagatcgaca	agctggtgag	caagggcatc	2340
cgcaagggtg	tgttcctgga	cggcatcgat	ggcggcacgc	tgatctacca	gtacatggac	2400
gacctgtacg	tgggcagcgg	cggccctagg	atcgattaaa	agcttcccgg	ggctagcacc	2460
ggtgaattc						2469

<210> 31

<211> 2463

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PR975YM

<400> 31

gtcgacgcca	ccatggccga	ggccatgagc	caggccacca	gcgccaacat	cctgatgcag	60
cgcagcaact	tcaagggccc	caagcgcata	atcaagtgt	tcaactgcgg	caaggagggc	120
cacatcgccc	gcaactgccg	cgccccccgc	aagaagggt	gctggaagt	cggaaggag	180
ggccaccaga	tgaaggactg	caccgagcgc	caggccaact	tcttccgcga	ggacctggcc	240
ttccccccag	gcaaggcccc	cgagttcccc	agcgagcaga	accgcgcaa	cagccccacc	300
agccgcgagc	tgcaggtgcg	cggcgacaac	ccccgcagcg	aggccggcgc	cgagcgccag	360
ggcaccctga	acttccccca	gatcacccct	tggcagcgcc	ccctggtgag	catcaaggtg	420
ggcggccaga	tcaaggaggc	cctgctggac	accggcgccg	acgacaccgt	gctggaggag	480
atgagcctgc	ccggcaagtg	gaagcccaag	atgatcggcg	gcatcggcgg	cttcatcaag	540
gtgcgccagt	acgaccagat	cctgatcgag	atctgcggca	agaaggccat	cggcaccgtg	600
ctgatcgccc	ccacccccgt	gaacatcatc	ggcgcgaaca	tgctgacca	gctgggctgc	660
accctgaact	tccccatcag	ccccatcgag	accgtgcccg	tgaagctgaa	gcccggcatg	720
gacgggcccc	aggtgaagca	gtggccccctg	accgaggaga	agatcaaggc	cctgaccgcc	780
atctgcgagg	agatggagaa	ggagggcaag	atcaccaaga	tcggccccga	gaacccctac	840
aacacccccg	tgttcgccat	caagaagaag	gacagcacca	agtggcgcaa	gctggtggac	900
ttccgcgagc	tgaacaagcg	caccacggac	ttctgggagg	tgcagctggg	catccccac	960
cccgccggcc	tgaagaagaa	gaagagcgtg	accgtgctgg	acgtgggcca	cgcctacttc	1020
agcgtgcccc	tggacgagga	cttccgcaag	tacaccgcct	tcaccatccc	cagcatcaac	1080
aacgagaccc	ccggcatccg	ctaccagtac	aacgtgctgc	cccagggctg	gaagggcagc	1140
cccagcatct	tccagagcag	catgaccaag	atcctggagc	ccttccgcgc	ccgcaacccc	1200
gagatcgtga	tctaccaggc	ccccctgtac	gtgggcagcg	acctggagat	cggccagcac	1260

cgcgccaaga	tcgaggagct	gcgcaagcac	ctgctgcgct	ggggcttcac	cacccccgac	1320
aagaagcacc	agaaggagcc	ccccttcctg	tggatgggct	acgagctgca	ccccgacaag	1380
tggaccgtgc	agcccatcga	gctgcccag	aaggagagct	ggaccgtgaa	cgacatccag	1440
aagctggtgg	gcaagctgaa	ctgggcccagc	cagatctacc	ccggcatcaa	ggtgcccag	1500
ctgtgcaagc	tgctgcgcgg	cgccaaggcc	ctgaccgaca	tcgtgcccct	gaccgaggag	1560
gccgagctgg	agctggccga	gaaccgcgag	atcctgcgcg	agcccgtgca	cggcgtgtac	1620
tacgacccca	gcaaggacct	ggtggccgag	atccagaagc	agggccacga	ccagtggacc	1680
taccagatct	accaggagcc	cttcaagaac	ctgaagaccg	gcaagtacgc	caagatgcgc	1740
accgcccaca	ccaacgacgt	gaagcagctg	accgaggccg	tgcagaagat	cgccatggag	1800
agcatcgtga	tctggggcaa	gacccccaa	ttccgcctgc	ccatccagaa	ggagacctgg	1860
gagacctggt	ggaccgacta	ctggcaggcc	acctggatcc	ccgagtggga	gttcgtgaac	1920
accccccccc	tggtgaagct	gtggtaccag	ctggagaagg	agcccatcat	cggcgccgag	1980
accttctacg	tggacggcgc	cgccaaccgc	gagaccaaga	tcggcaaggc	cggctacgtg	2040
accgaccggg	gccgctagaa	gatcgtgagc	ctgaccgaga	ccaccaacca	gaagaccgag	2100
ctgcaggcca	tccagctggc	cctgcaggac	agcggcagcg	aggtgaacat	cgtgaccgac	2160
agccagtacg	ccctgggcat	catccaggcc	cagcccagaca	agagcgagag	cgagctggtg	2220
aaccagatca	tcgagcagct	gatcaagaag	gagaagggtg	acctgagctg	ggtgcccgcc	2280
cacaagggca	tcggcgga	cgagcagatc	gacaagctgg	tgagcaagg	catccgcaag	2340
gtgctgttcc	tggacggcat	cgatggcggc	atcgtgatct	accagtacat	ggacgacctg	2400
tacgtgggca	gcggcgccc	taggatcgat	taaaagcttc	ccggggctag	caccggtgaa	2460
ttc						2463

<210> 32

<211> 2457

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PR975YMWM

<400> 32

gtcgacgcca	ccatggccga	ggccatgagc	caggccacca	gcgccaacat	cctgatgcag	60
cgagcaact	tcaaggccc	caagcgcac	atcaagtgc	tcaactgcg	caaggaggc	120
cacatcgccc	gcaactgcc	cgccccccg	aagaagggt	gctggaagt	cggcaaggag	180
ggccaccaga	tgaaggact	caccgagcgc	caggccaact	tcttcgcga	ggacctggc	240
ttcccccagg	gcaaggccc	cgagttcccc	agcagcaga	accgcgcaa	cagccccacc	300
agccgcgagc	tgcaagtg	cggcgacaac	ccccgcagc	aggccggcgc	cgagcgccag	360
ggcaccctga	acttccccca	gatcaccctg	tggcagcgcc	ccctggtgag	catcaagggtg	420
ggcggccaga	tcaaggaggc	cctgctggac	accggcgccg	acgacaccgt	gctggaggag	480
atgagcctgc	ccggcaagt	gaagcccaag	atgatcggc	gcacggcg	cttcatcaag	540
gtgcgccagt	acgaccagat	cctgatcgag	atctgcggca	agaaggccat	cggcaccgtg	600
ctgatcggcc	ccacccccgt	gaacatcatc	ggccgcaaca	tgctgacca	gctgggctgc	660
accctgaact	tccccatcag	ccccatcgag	accgtgccc	tgaagctgaa	gcccggcatg	720
gacggcccca	aggtgaagca	gtggcccctg	accgaggaga	agatcaaggc	cctgaccgcc	780
atctgcgagg	agatggagaa	ggagggcaa	atcaccaaga	tcggccccga	gaacccctac	840
aacacccccg	tgctcgccat	caagaagaag	gacagcacca	agtggcgcaa	gctggtggac	900
ttccgcgagc	tgaacaagcg	caccaggac	ttctgggagg	tgacgtggg	catccccac	960
cccgcgggcc	tgaagaagaa	gaagagcgtg	accatgctgg	acgtgggcga	cgcctacttc	1020
agcgtgcccc	tggacgagga	cttcgcgaag	tacacgcct	tcacatccc	cagcatcaac	1080
aacgagaccc	ccggcatccg	ctaccagtat	aacgtgctgc	cccagggtg	gaaggcgagc	1140
cccagcatct	tccagagcag	catgaccaag	atcctggagc	ccttcgcgc	ccgcaacccc	1200
gagatcgtga	tctaccaggc	ccccctgtac	gtgggcagcg	acctggagat	cggccagcac	1260
cgcgccaaga	tcgaggagct	gcgcaagcac	ctgctgcgct	ggggcttcac	cacccccgac	1320
aagaagcacc	agaaggagcc	ccccttcctg	cccatcgagc	tgcaccccg	caagtggacc	1380
gtgcagccca	tcgagctgcc	cgagaaggag	agctggaccg	tgaacgacat	ccagaagctg	1440
gtgggcaagc	tgaactgggc	cagccagatc	taccccgga	tcaaggtgcg	ccagctgtgc	1500
aagctgctgc	gcggcgccaa	ggccctgacc	gacatcgtgc	ccctgaccga	ggaggccgag	1560

ctggagctgg	ccgagaaccg	cgagatcctg	cgcgagcccc	tgcacggcgt	gtactacgac	1620
cccagcaagg	acctggtggc	cgagatccag	aagcagggcc	acgaccagtg	gacctaccag	1680
atctaccagg	agcccttcaa	gaacctgaag	accggcaagt	acgccaagat	gcgaccgcc	1740
cacaccaacg	acgtgaagca	gctgaccgag	gccgtgcaga	agatcgccat	ggagagcatc	1800
gtgatctggg	gcaagacccc	caagttccgc	ctgccccatcc	agaaggagac	ctgggagacc	1860
tgggtggaccg	actactggca	ggccacctgg	atccccgagt	gggagttcgt	gaacaccccc	1920
cccctggtga	agctgtggta	ccagctggag	aaggagccca	tcatcggcgc	cgagaccttc	1980
tacgtggacg	gcgccgcca	ccgcgagacc	aagatcggca	aggccggcta	cgtgaccgac	2040
cggggccggc	agaagatcgt	gagcctgacc	gagaccacca	accagaagac	cgagctgcag	2100
gccatccagc	tggccctgca	ggacagcggc	agcgaggtga	acatcgtgac	cgacagccag	2160
tacgccttgg	gcatcatcca	ggcccagccc	gacaagagcg	agagcgagct	ggtgaaccag	2220
atcatcgagc	agctgatcaa	gaaggagaag	gtgtacctga	gctgggtgcc	cgcccacaag	2280
ggcatcggcg	gcaacgagca	gatcgacaag	ctggtgagca	agggcatccg	caaggtgctg	2340
ttcctggacg	gcatcgatgg	cggcatcgtg	atctaccagt	acatggacga	cctgtactgtg	2400
ggcagcggcg	gccctaggat	cgattaaaag	cttcccgggg	ctagcaccgg	tgaattc	2457

<210> 33

<211> 9781

<212> DNA

<213> Human immunodeficiency virus

<400> 33

tggaaggggt	aatttactcc	aagaaaaggc	aagaaatcct	tgatttgtgg	gtctatcaca	60
cacaaggctt	cttccctgat	tggcaaaact	acacaccggg	gccaggggtc	agatatccac	120
tgacctttgg	atggtgctac	aagctagtgc	cagttgacct	aggggaggtg	gaagaggcca	180
acggaggaga	agacaactgt	ttgctacacc	ctatgagcca	acatggagca	gaggatgaag	240
atagagaagt	attaaagtgg	aagtttgaca	gcctcctagc	acgcagacac	atggcccgcg	300
agctacatcc	ggagtattac	aaagactgct	gacacagaag	ggactttccg	cctgggactt	360
tccactgggg	cgttccggga	ggtgtggtct	gggcgggact	tgggagtgg	caaccctcag	420
atgctgcata	taagcagctg	cttttcgcct	gtactgggtc	tctctcggtg	gaccagatct	480
gagcctggga	gccctctggc	tatctaggga	accactgct	taagcctcaa	taaagcttgc	540
cttgagtgtc	ttaagtagtg	tgtgcccata	tgttggtgta	ctctggtaac	tagagatccc	600
tcagaccctt	tgtggtagt	tggaaaatct	ctagcagtgg	cgcccgaaca	gggaccagaa	660
agtgaagt	agaccagagg	agatctctcg	acgcaggact	cggcttgctg	aagtgcacac	720
ggcaagaggc	gagagggg	gctggtgagt	acgccaattt	tacttgacta	gcggaggcta	780
gaaggagaga	gatgggtgcg	agagcgtcaa	tattaagcgg	cggaaaatta	gataaatggg	840
aaagaattag	gttaaggcca	gggggaaaga	aacattatat	gttaaaacat	ctagtatggg	900
caagcaggga	gctggaaaga	tttgcaactta	accctggcct	gttagaaaaca	tcagaaggct	960
gtaaacaat	aataaaacag	ctacaaccag	ctcttcagac	aggaacagag	gaacttagat	1020
cattattcaa	cacagtagca	actctctatt	gtgtacataa	agggatagag	gtacgagaca	1080
ccaaggaagc	cttagacaag	atagaggaag	aacaaaacaa	atgtcagcaa	aaagcacaac	1140
aggcaaaagc	agctgacgaa	aaggtcagtc	aaaattatcc	tatagtacag	aatgcccaag	1200
ggcaaatggt	acaccaagct	atatcaccta	gaacattgaa	tgcatggata	aaagtaatag	1260
aggaaaaggc	tttcaatcca	gaggaaatac	ccatgtttac	agcattatca	gaaggagcca	1320
cccacaaga	tttaaacaca	atgttaaata	cagtgggggg	acatcaagca	gccatgcaaa	1380
tgttaaaaga	taccatcaat	gaggaggctg	cagaatggga	taggacacat	ccagtacatg	1440
cagggcctgt	tgaccaggc	cagatgagag	aaccaagggg	aagtgcacata	gcaggaacta	1500
ctagtaccct	tcaggaacaa	atagcatgga	tgacaagtaa	tccacctatt	ccagtagaag	1560
acatctataa	aagatggata	attctggggt	taataaaaat	agtaagaatg	tatagccctg	1620
ttagcatttt	ggacataaaa	caagggccaa	aagaaccctt	tagagactat	gtagaccgtg	1680
tctttaaaac	cttaagagct	gaacaagcta	cacaagatgt	aaagaattgg	atgacagaca	1740
ccttgttggt	ccaaaatgcg	aaccagatt	gtaagaccat	tttaagagca	ttaggaccag	1800
gggcctcatt	agaagaaatg	atgacagcat	gtcagggagt	gggaggacct	agccataaag	1860
caagagtgtt	ggctgaggca	atgagccaag	caaacagtaa	catactagt	cagagaagca	1920
attttaaagg	ctctaacaga	attattaaat	gtttcaactg	tggcaaagta	gggcacatag	1980
ccagaaattg	cagggccct	aggaaaaagg	gctgttgga	atgtggacag	gaaggacacc	2040
aatgaaaga	ctgtactgag	aggcaggcta	attttttagg	gaaaatttgg	ccttcccaca	2100

aggggaggcc	aggggaatttc	ctccagaaca	gaccagagcc	aacagcccca	ccagcagaac	2160
caacagcccc	accagcagag	agcttcaggt	tcgaggagac	aacccccgtg	ccgaggaagg	2220
agaaagagag	ggaaccttta	acttcctca	aatcactctt	tggcagcgac	cccttgctc	2280
aataaaagta	gagggccaga	taaaggaggc	tctcttagac	acaggagcag	atgatacagt	2340
attagaagaa	atagatttgc	cagggaaatg	gaaacccaaa	atgatagggg	gaattggagg	2400
ttttatcaaa	gtaagacagt	atgatcaaat	acttatagaa	atgttgagg	aaaaggctat	2460
aggtacagta	ttagtagggc	ctacaccagt	caacataatt	ggaagaaatc	tgtaaactca	2520
gcttgatgc	acactaaatt	ttccaattag	tcctattgaa	actgtaccag	taaaattaaa	2580
accaggaatg	gatggcccaa	aggtcaaaca	atggccattg	acagaagaaa	aaataaaagc	2640
attaacagca	atgttgtagg	aaatggagaa	ggaaggaaaa	attacaaaaa	ttgggcctga	2700
taatccatat	aacactccag	tatttgccat	aaaaaagaag	gacagtacta	agtggagaaa	2760
attagtagat	ttcaggggaac	tcaataaaaag	aactcaagac	ttttgggaag	ttcaattagg	2820
aataccacac	tcagcaggat	taaaaaagaa	aaaatcagtg	acagtgcctag	atgtggggga	2880
tgcatatttt	tcagttcctt	tagatgaaag	cttcaggaaa	tatactgcat	tcaccatacc	2940
tagtataaac	aatgaaacac	cagggattag	atatcaatat	aatgtgctgc	cacagggatg	3000
gaaaggatca	ccagcaatat	tccagagtag	catgacaaaa	atcttagagc	ccttcagagc	3060
aaaaaatcca	gacatagtta	tctatcaata	tatggatgac	ttgtatgtag	gatctgactt	3120
agaaataggg	caacatagag	caaaaataga	agagttaagg	gaacatttat	tgaaatgggg	3180
atttacaaca	ccagacaaga	aacatcaaaa	agaaccccca	tttctttgga	tgggggatga	3240
actccatcct	gacaaatgga	cagtacaacc	tatactgctg	ccagaaaagg	atagttggac	3300
tgtcaatgat	atacagaagt	tagtgggaaa	attaaactgg	gcaagtcaga	tttaccagg	3360
gattaaagta	aggcaactct	gtaaaactcct	cagggggggc	aaagcactaa	cagacatagt	3420
accactaact	gaagaagcag	aattagaatt	ggcagagaac	agggaaattt	taagagaacc	3480
agtacatgga	gtatattatg	atccatcaaa	agacttgata	gctgaaatac	agaaacaggg	3540
gcatgaacaa	tggacatatc	aaatttatca	agaaccattt	aaaaatctga	aaacagggga	3600
gtatgcaaaa	atgaggacta	cccacactaa	tgatgtaaaa	cagttaacag	aggcagtgca	3660
aaaaatagcc	atggaaagca	tagtaatatg	gggaaagact	cctaaattta	gactacccat	3720
ccaaaaagaa	acatgggaga	catggtggac	agactattgg	caagccacct	ggatccctga	3780
gtgggagttt	gttaataccc	ctcccctagt	aaaattatgg	taccaactag	aaaaagatcc	3840
catagcagga	gtagaaactt	tctatgtaga	tggagcaact	aatagggaag	ctaaaatagg	3900
aaaagcaggg	tatgttactg	acagaggaag	gcagaaat	gttactctaa	ctaacacaac	3960
aaatcagaag	actgagttac	aagcaattca	gctagctctg	caggattcag	gatcagaagt	4020
aaacatagta	acagactcac	agtatgcatt	aggaatcatt	caagcacaa	cagataagag	4080
tgactcagag	atatttaacc	aaataataga	acagttaata	aacaaggaaa	gaatctacct	4140
gtcatgggta	ccagcacata	aaggaattgg	gggaaatgaa	caagtagata	aattagtaag	4200
taagggaatt	aggaaagtgt	tgtttctaga	tggaaatagat	aaagctcaag	aagagcatga	4260
aaggtaccac	agcaattgga	gagcaatggc	taatgagttt	aatctgccac	ccatagtagc	4320
aaaagaaata	gtagctagct	gtgataaatg	tcagctaaaa	ggggaagcca	tacatggaca	4380
agtcgactgt	agtccaggga	tatggcaatt	agattgtacc	catttagagg	gaaaaatcat	4440
cctggtagca	gtccatgtag	ctagtggcta	catggaagca	gaggttatcc	cagcagaaac	4500
aggacaagaa	acagcatatt	ttatattaaa	attagcagga	agatggccag	tcaaagtaat	4560
acatacagac	aatggcagta	attttaccag	tactgcagtt	aaggcagcct	gttggtgggc	4620
aggtatccaa	caggaatttg	gaattcccta	caatcccaa	agtcaggagg	tggtagaatc	4680
catgaataaa	gaattaaaga	aaataatagg	acaagtaaga	gatcaagctg	agcaccttaa	4740
gacagcagta	caaatggcag	tattcattca	caattttaaa	agaaaagggg	gaattggggg	4800
gtacagtgca	ggggaagaa	taatagacat	aatagcaaca	gacatacaaa	ctaaagaatt	4860
acaaaaacaa	attataagaa	ttcaaaattt	tcgggtttat	tacagagaca	gcagagaccc	4920
tatttggaag	ggaccagccg	aactactctg	gaaaggtgaa	ggggtagtag	taatagaaga	4980
taaagggtgac	ataaaggtag	taccaaggag	gaaagcaaaa	atcattagag	attatggaaa	5040
acagatggca	ggtgctgatt	gtgtggcagg	tggacaggat	gaagattaga	gcatggaata	5100
gttttagtaaa	gcaccatatg	tatatatcaa	ggagagctag	tggatgggtc	tacagacatc	5160
attttgaaag	cagacatcca	aaagtaagtt	cagaagtaca	tatccatta	ggggatgcta	5220
gattagtaat	aaaaacatat	tggggtttgc	agacaggaga	aagagattgg	catttgggtc	5280
atggagtctc	catagaatgg	agactgagag	aatacagcac	acaagtagac	cctgacctgg	5340
cagaccagct	aattcacatg	cattattttg	attgttttac	agaatctgcc	ataagacaag	5400
ccatattagg	acacatagtt	tttcctaggt	gtgactatca	agcaggacat	aagaaggtag	5460
gatctctgca	atacttggca	ctgacagcat	tgataaaacc	aaaaaagaga	aagccacctc	5520

tgcctagtgt	tagaaaatta	gtagaggata	gatggaacga	ccccagaag	accagggggcc	5580
gcagagggaa	ccatacaatg	aatggacact	agagattcta	gaagaactca	agcaggaagc	5640
tgtcagacac	tttcctagac	catggctcca	tagcttagga	caatatactt	atgaaaccta	5700
tggggatact	tggacgggag	ttgaagctat	aataagagta	ctgcaacaac	tactgttcat	5760
tcattttcaga	attggatgcc	aacatagcag	aataggcatc	ttgcgacaga	gaagagcaag	5820
aaatggagcc	agtagatcct	aaactaaagc	cctggaacca	tccaggaagc	caacctaaaa	5880
cagcttgtaa	taattgcttt	tgcaaact	gtagctatca	ttgtctagtt	tgctttcaga	5940
caaaaggttt	aggcatttcc	tatggcagga	agaagcggag	acagcgacga	agcgtctctc	6000
caagtgggtga	agatcatcaa	aatcctctat	caaagcagta	agtacacata	gtagatgtaa	6060
tggtaaagttt	aagtttattt	aaaggagtag	attatagatt	aggagtagga	gcattgatag	6120
tagcactaat	catagcaata	atagtgtgga	ccatagcata	tatagaatat	aggaaattgg	6180
taagacaaaa	gaaaatagac	tggtttaatta	aaagaattag	ggaaagagca	gaagacagtg	6240
gcaatgagag	tgatggggag	acagaagaat	tgtcaacaat	ggtggatatg	gggcatctta	6300
ggcttctgga	tgctaattgat	ttgtaaacag	gaggacttgt	gggtcacagt	ctactatggg	6360
gtacctgtgt	ggagagaagc	aaaaactact	ctattctgtg	catcagatgc	taaagcatat	6420
gagacagaag	tgcataatgt	ctgggctaca	catgcttgtg	taccacaga	ccccaaccca	6480
caagaaatag	ttttgggaaa	tgtaacagaa	aatttttaata	tgtggaaaaa	taacatggca	6540
gatcagatgc	atgaggatat	aatcagttta	tgggatcaaa	gcctaaagcc	atgtgtaaag	6600
ttgacccac	tctgtgtcac	tttaaactgt	acagatacaa	atgttacagg	taatagaact	6660
gttacaggtg	atacaaatga	taccaatatt	gcaaattgcta	catataagta	tgaagaaatg	6720
aaaaattgct	ctttcaatgc	aaccacagaa	ttaagagata	agaaacataa	agagtatgca	6780
ctcttttata	aacttgatat	agtaccactt	aatgaaaata	gtaacaactt	tacatataga	6840
ttaataaatt	gcaatacctc	aaccataaca	caagcctgtc	caaaggtctc	ttttgacctg	6900
attcctatac	attactgtgc	tccagctgat	tatgcgattc	taaagtgtaa	taataagaca	6960
ttcaatggga	caggaccatg	ttataatgtc	agcacagtac	aatgtacaca	tggaattaaag	7020
ccagtgggtat	caactcaact	actgtttaat	ggtagtctag	cagaagaagg	gataataatt	7080
agatctgaaa	atttgacaga	gaataccaaa	acaataatag	tacatcttaa	tgaatctgta	7140
gagattaatt	gtacaaggcc	caacaataat	acaaggaaaa	gtgtaaggat	aggaccagga	7200
caagcattct	atgcaacaaa	tgacgtaata	ggaaacataa	gacaagcaca	ttgtaacatt	7260
agtacagata	catggaataa	aactttacaa	caggtaatga	aaaaattagg	agagcatttc	7320
cctaataaaa	caataaaatt	tgaaccacat	gcaggagggg	atctagaaat	tacaatgcct	7380
agctttaatt	gtagaggaga	atTTTTctat	tgcaatacat	caaacctgtt	taatagtaca	7440
tactacccta	agaatggtac	atacaaatat	aatggtaatt	caagcttacc	catcacactc	7500
caatgcaaaa	taaaacaaat	tgtacgcatg	tggcaagggg	taggacaagc	aatgtatgcc	7560
cctcccatgt	caggaaacat	aacatgtaga	tcaaactca	caggaaatac	attgacacgt	7620
gatgggggat	ttacaacac	aaacaacgac	acagaggaga	cattcagacc	tggaggagga	7680
gatatgaggg	ataactggag	aagtgaatta	tataaatata	aagtggtaga	aattaagcca	7740
ttgggaatag	cacccactaa	ggcaaaaaga	agagtgggtc	agagaaaaaa	aagagcagtg	7800
ggaataggag	ctgtgttcct	tgggttcttg	ggagcagcag	gaagcactat	gggcgcagcg	7860
tcaataacgc	tgacggtaca	ggccagacaa	ctgttgtctg	gtatagtgc	acagcaaagc	7920
aatttgctga	aggctataga	ggcgcaacag	catatgttgc	aactcacagt	ctggggcatt	7980
aagcagctcc	aggcgagagt	cctggctata	gaaagatacc	taaaggatca	acagctccta	8040
gggatttggg	gctgctctgg	aagactcatc	tgcaccactg	ctgtgccttg	gaactccagt	8100
tgtagtaata	aatctgaagc	agatatTTtg	gataacatga	cttggatgca	gtgggataga	8160
gaaattaata	attacacaga	aacaatatct	aggttgcttg	aagactcgca	aaaccagcag	8220
gaaaagaatg	aaaaagattt	attagaattg	gacaagtgga	ataatctgtg	gaattggttt	8280
gacatatcaa	actggctgtg	gtatataaaa	atattcataa	tgatagtagg	aggcttgata	8340
ggtttaagaa	taatttttgc	tgtgctctct	atagtgaata	gagttaggca	gggatactca	8400
cctttgtcat	ttcagaccct	tacccaagc	ccgaggggac	tcgacaggct	cggaggaatc	8460
gaagaagaag	gtggagagca	agacagagac	agatccatac	gattgggtgag	cggattcttg	8520
tcgcttgctt	gggacgatct	gcggagcctg	tgcctcttca	gctaccaccg	cttgagagac	8580
ttcatattaa	ttgcagttag	ggcagtggaa	cttctgggac	acagcagtct	caggggacta	8640
cagagggggg	gggagatcct	taagtatctg	ggaagtcttg	tgcatgattg	gggtctagag	8700
ctaaaaaaga	gtgctattag	tccgcttgat	accatagcaa	tagcagtagc	tgaaggaaca	8760
gataggatta	tagaattggg	acaaagaatt	tgtagagcta	tcctcaacat	acctaggaga	8820
ataagacagg	gctttgaagc	agctttgcta	taaaatggga	ggcaagtggg	caaaacgcag	8880
catagttgga	tggcctgcag	taagagaaa	aatgagaaga	actgagccag	cagcagaggg	8940

agtaggagca	gcgtctcaag	acttagatag	acatggggca	cttacaagca	gcaacacacc	9000
tgctactaat	gaagcttgtg	cctggctgca	agcacaagag	gaggacggag	atgtaggctt	9060
tccagtcaga	cctcaggtac	ctttaagacc	aatgacttat	aagagtgcag	tagatctcag	9120
cttcttttta	aaagaaaagg	ggggactgga	aggggttaatt	tactctagga	aaaggcaaga	9180
aatccttgat	ttgtgggtct	ataacacaca	aggcttcttc	cctgattggc	aaaactacac	9240
atcggggcca	ggggtccgat	tcccactgac	ctttggatgg	tgcttcaagc	tagtaccagt	9300
tgaccaaggg	gaggtgaaag	aggccaatga	aggagaagac	aactgtttgc	tacaccctat	9360
gagccaacat	ggagcagagg	atgaagatag	agaagtatta	aagtgggaagt	ttgacagcct	9420
tctagcacac	agacacatgg	cccgcgagct	acatccggag	tattacaaag	actgctgaca	9480
cagaagggac	tttccgcctg	ggactttcca	ctggggcggt	ccgggaggtg	tggtctgggc	9540
gggacttggg	agtgttcacc	ctcagatgct	gcatataagc	agctgctttt	cgcttgact	9600
gggtctctct	cggtagacca	gatctgagcc	tgggagctct	ctggctatct	agggaacca	9660
ctgcttaggc	ctcaataaag	cttgccctga	gtgctctaag	tagtgtgtgc	ccatctgttg	9720
tgtagctctg	gtaactagag	atccctcaga	ccctttgtgg	tagtgtggaa	aatctctagc	9780
a						9781

<210> 34

<211> 203

<212> DNA

<213> Human immunodeficiency virus

<400> 34

gctgaggcaa	tgagccaagc	aaccagcgca	aacatactga	tgagagaag	caatttcaaa	60
ggccctaaaa	gaattattaa	atgtttcaac	tgtggcaagg	aagggcacat	agctagaaat	120
tgtagggccc	ctaggaaaaa	aggctgttgg	aaatgtggaa	aggaaggaca	ccaaatgaaa	180
gactgtactg	agaggcaggc	taa				203

<210> 35

<211> 2151

<212> DNA

<213> Human immunodeficiency virus

<400> 35

tttttttaggg	aagatttggc	cttcccacaa	gggaaggcca	gggaatttcc	ttcagaacag	60
aacagagcca	acagccccac	cagcagagag	cttcaagttc	gaggagacaa	ccccgctcc	120
gaagcaggag	ccgaaagaca	gggaaccctt	aatttccctc	aaatcactct	ttggcagcga	180
ccccttgtct	caataaaagt	aggggggtcaa	ataaaggagg	ctctcttaga	cacaggagct	240
gatgatacag	tattagaaga	aatgagtttg	ccaggaaaat	ggaaaccaa	aatgatagga	300
ggaattggag	gttttatcaa	agtaagacag	tatgatcaaa	tacttataga	aatttgtgga	360
aaaaaggcta	taggtacagt	attaatagga	cctacacctg	tcaacataat	tggaagggaat	420
atgttgactc	agcttggatg	cacactaaat	tttccaatta	gtcccattga	aactgtgcc	480
gtaaaattaa	agccaggaat	ggatggccca	aagggttaa	aatggccatt	gacagaagag	540
aaaataaaaag	cattaacagc	aatttgtgaa	gaaatggaga	aagaaggaaa	aattacaaaa	600
attgggcctg	aaaatccata	taacactcca	gtatttggca	taaaaaagaa	ggacagtact	660
aagtggagaa	agttagtaga	tttcagggaa	cttaataaaa	gaactcaaga	cttttgggaa	720
gttcaattag	gaataccaca	cccagcaggg	ttaaaaaaga	aaaaatcagt	gacagtactg	780
gatgtggggg	atgcatattt	ttcagttcct	ttagatgagg	acttcaggaa	atatactgca	840
ttcaccatac	ctagtataaa	caatgaaaca	ccagggatta	gatatcaata	taatgtgctt	900
ccacagggat	ggaaaggatc	accatcaata	ttccagagta	gcatgacaaa	aatcttagag	960
cccttttagag	caagaaatcc	agaaatagtc	atctatcaat	atatggatga	cttgtatgta	1020
ggatctgact	tagaaatagg	gcaacataga	gcaaaaaatag	aggagttaa	aaaacatctg	1080
ttaaggtggg	gatttaccac	accggacaag	aaacatcaga	aagaaccccc	atttcttttg	1140
atgggggatg	aactccatcc	tgacaaatgg	acagtacagc	ctatagagtt	gccagaaaag	1200
gaaagctgga	ctgtcaatga	tatacagaag	ttagtgggaa	aattaaattg	ggccagtcag	1260
atttaccagg	gaattaaagt	aaggcaactt	tgtaaactcc	ttaggggggc	caaagcacta	1320
acagatatag	taccactaac	tgaagaagca	gaattagaat	tggcagagaa	cagggaaatt	1380
ctaagagaac	cagtacatgg	agtatattat	gacccatcaa	aagacttggg	agctgaaata	1440

```

cagaaacagg ggcatgacca atggacatat caaatattacc aagaaccatt caaaaacctg 1500
aaaacaggga agtatgcaaa aatgaggact gccacacta atgatgtaaa acagttaaca 1560
gaggcagtg caaaaaatagc tatggaaagc atagtaatat ggggaaagac tcctaaattt 1620
agactacca tccaaaaaga aacatgggag acatggtgga cagactattg gcaagccacc 1680
tggttcctg agtgggagtt tgtaataacc cctcccttag taaaattatg gtaccagcta 1740
gagaaagaac ccataatagg agcagaaact ttctatgtag atggagcagc taatagggaa 1800
actaaaatag gaaaagcagg gtatgttact gacagaggaa ggcagaaaat tgtttctcta 1860
acagaaacaa caaatcagaa gactgaatta caagcaattc agctagcttt gcaagattca 1920
ggatcagaag taaacatagt aacagactca cagtatgcat taggaatcat tcaagcacia 1980
ccagataaga gtgaatcaga gttagtcaac caaataatag aacaattaat aaaaaaggaa 2040
aaggtctacc tgtcatgggt accagcacat aaaggaattg gaggaatga acaaatagat 2100
aaattagtaa gtaagggaat caggaaagtg ctgtttctag atggaataga t 2151

```

<210> 36
 <211> 54
 <212> DNA
 <213> Human immunodeficiency virus

```

<400> 36
ggcggcatcg tgatctacca gtacatggac gacctgtacg tgggcagcgg cggc 54

```

<210> 37
 <211> 18
 <212> PRT
 <213> Human immunodeficiency virus

```

<400> 37
Gly Gly Ile Val Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val Gly Ser
  1             5             10             15

```

Gly Gly

<210> 38
 <211> 38
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer
 S1FCSacTA

```

<400> 38
gtttcttgag ctctggaagg gttaatttac tccaagaa 38

```

<210> 39
 <211> 38
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence: primer
 S1FTSactA

```

<400> 39
gtttcttgag ctctggaagg gttaatttac tctaagaa 38

```

<210> 40
 <211> 35
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: primer
 S145RTSalTA

 <400> 40
 gtttcttgtc gacttggtcca tgtatggctt cccct 35

 <210> 41
 <211> 34
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: primer
 S145RCSalTA

 <400> 41
 gtttcttgtc gacttggtcca tgcattggctt ccct 34

 <210> 42
 <211> 38
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: primer
 S245FASalTA

 <400> 42
 gtttcttgtc gactgtagtc caggaatatg gcaattag 38

 <210> 43
 <211> 38
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: primer
 S245FGSalTA

 <400> 43
 gtttcttgtc gactgtagtc cagggatatg gcaattag 38

 <210> 44
 <211> 39
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> Description of Artificial Sequence: primer
 S2FullNotTA

<400> 44
gtttcttgcg gccgctgcta gagattttcc acactacca

39

<210> 45
<211> 9738
<212> DNA
<213> Human immunodeficiency virus

<400> 45
tggaagggtt aatttactcc aggaaaaggc aagagatcct tgatttatgg gtctatcaca 60
cacaaggcta cttccctgat tggcaaaact acacaccggg accaggggtc agatatccac 120
tgacctttgg atggtgcttc aagctagtgc cagttgacct aagggaagta gaagaggcca 180
acggaggaga agacaactgt ttgctacacc ctatgagcca gtatggaatg gatgatgaac 240
acaaagaagt gttacagtgg aagtttgaca gcagcctagc acgcagacac ctggcccgcg 300
agctacatcc ggattattac aaagactgct gacacagaag ggactttccg cctgggactt 360
tccactgggg cgttccaggg ggagtggctt gggcgggact gggagtggcc agccctcaga 420
tgctgcatat aagcagcggc ttttcgcctg tactgggtct ctctaggtag accagatccg 480
agcctgggag ctctctgtct atctggggaa cccactgctt aggcctcaat aaagcttgcc 540
ttgagtgtct taagtagtgt gtgcccctct gttgtgtgac tctggttaact ctggttaacta 600
gagatccctc agaccctttg tggtagtgtg gaaaatctct agcagtggcg cccgaacagg 660
gacttgaaag cgaaagtgag accagagaag atctctcgac gcaggactcg gcttgctgaa 720
gtgcactcgg caagaggcga ggggggcgac tggtagtagt gccaaaattt tttttgacta 780
gcggaggcta gaaggagaga gatgggtgcg agagcgtcaa tattaagagg gggaaaatta 840
gacaaatggg aaaaaattag gttacggcca ggggggagaa aacactatat gctaaaacac 900
ctagtatggg caagcagaga gctggaaaga tttgcagtta accctggcct tttagagaca 960
tcagacggat gtagacaaat aataaaacag ctacaaccag ctcttcagac aggaacagag 1020
gaaattagat cattatttaa cacagtagca actctctatt gtgtacataa agggatagat 1080
gtacgagaca ccaaggaagc cttagacaag atagaggagg aacaaaacaa atgtcagcaa 1140
aaaacacagc aggcggaagc ggctgacaaa aaggtcagtc aaaattatcc tatagtgcag 1200
aacctccaag ggcaaatggt acaccaggcc atatcaccta gaaccttgaa tgcatgggta 1260
aaagtaatag aggagaaggc ttttagccca gaggtaatac ccatgtttac agcattatca 1320
gaaggagcca cccacaaga tttaaacacc atgttaaata cagtgggggg acatcaagca 1380
gccatgcaaa tgttaaaaga taccatcaat gaggaggctg cagaatggga taggttacat 1440
ccagtacatg cagggcctgt tgcaccaggc cagatgagag aaccaagggg aagtgcata 1500
gcaggaacta ctagtaccct tcaagaacaa atagcatgga tgacaagtaa cccacctatc 1560
ccagtagggg acatctataa aaggtggata attctggggg taaataaaat agtaagaatg 1620
tacagccctg tcagcatttt agacataaaa caaggaccaa aggaaccctt tagagactat 1680
gtagaccggt tcttcaaaac ttttaagagc gaacaatcta cacaagaggt aaaaaattgg 1740
atgacagaca ccttgttagt ccaaaatgcg aaccagatt gtaagaccat ttttaagagca 1800
ttaggaccag gggcttcatt agaagaaatg atgacagcat gtcagggagt gggaggacct 1860
agccacaaag caagagtttt ggctgaggca atgagccaag caaacaatac aagtgtaatg 1920
atacagaaaa gcaattttta aggccctaga agagctgtta aatgtttcaa ctgtggcagg 1980
gaagggcaca tagccaggaa ttgcagggcc cctaggaaaa ggggctgttg gaaatgtgga 2040
aagggaaggac accaaatgaa agactgtact gagaggcagg ctaatttttt agggaaaatt 2100
tgcccttccc acaaggggag gccagggaat ttccctcaga gcagaccaga gccaacagcc 2160
ccaccactag aaccaacagc cccaccagca gacagcttca agttcaagga gactccgaag 2220
caggagccga aagacaggga acctttaact tccctcaaat cactctttgg cagcgacccc 2280
ttgtctcaat aaaagtagcg ggccaaacaa aggaggctct tttagataga ggagcagatg 2340
atacagtact agaagaaata aacttgccag gaaaaatgaa accaaaaatg ataggaggaa 2400
ttggaggttt tatcaaagta agacagtatg atcaaatact tatagaaatt tgtggaaaaa 2460
gggctatagg tacagtatta gtaggacctt cacctgtcaa cataattgga agaaatctgt 2520
tgactcagct tggatgcaca ctaaaatttc caattagccc cattgaaact gtaccagtaa 2580
aattaaagcc aggaatggat ggcccaaagg ttaaacaatg gccattgaca gaagaaaaaa 2640
taaaagcatt aacagaaatt tgtgaggaaa tggagaagga aggaaaaatt acaaaaaattg 2700
ggcctgaaaa tccatataac actccagtat ttgccataaa gaagaaggac agtacaagat 2760

ggagaaaatt	agtagatttc	agggaactca	ataaaagaac	tcaagacttt	tgggaagtcc	2820
aattaggaat	accacaccca	gcagggttaa	aaaagaaaaa	atcagtgaca	gtactggatg	2880
tgggagatgc	atatttttca	gtccctttag	atgagagctt	cagaaaatat	actgcattca	2940
ccatacctag	tataaacaat	gaaacaccag	ggattagata	tcaatataat	gttcttccac	3000
agggatggaa	aggatcacca	gcaatattcc	agagtagcat	gacaagaatc	ttagagccct	3060
ttagaacaca	aaaccagaa	gtagtattct	atcaatatat	ggatgactta	tatgtaggat	3120
ctgacttaga	aatagggcaa	catagagcaa	aaatagagga	gttaagagga	cacctattga	3180
aatggggatt	taccacacca	gacaagaaac	atcagaaaga	acccccattt	ctttggatgg	3240
ggtatgaact	ccatcctgac	aaatggacag	tacagcctat	acagctgcca	gaaaaggaga	3300
gctggactgt	caatgatata	cagaagttag	tgggaaagtt	aaactgggca	agtcagattt	3360
acccagggat	taaagtaagg	caactgtgta	aactccttag	gggagccaaa	gcactaacag	3420
acatagtgcc	actgactgaa	gaagcagaat	tagaattggc	tgagaacagg	gaaattctaa	3480
aagaaccagt	acatggagta	tattatgacc	catcaaaaaga	tttaatagct	gaaatacaga	3540
aacaggggaa	tgaccaatgg	acatatcaaa	tttaccaga	accattttaa	aatctgagaa	3600
caggaaaagta	tgcaaaaatg	aggactgccc	acactaatga	tgtgaaacag	ttagcagagg	3660
cagtgcacaaa	gataacccag	gaaagcatag	taatattggg	aaaaactcct	aaatttagac	3720
tacccatccc	aaaagaaaca	tgggagacat	ggtggtcaga	ctattggcaa	gccacctgga	3780
ttcctgagtg	ggagtttgtc	aatacccttc	ccctagtaaa	attgtggtac	cagctggaaa	3840
aagaacccat	agtaggggca	gaaactttct	atgtagatgg	agcagccaat	agggaaacta	3900
aaataggaaa	agcaggggat	gtcactgaca	aaggaaggca	gaaagtgtgt	tccttcactg	3960
aaacaacaaa	tcagaagact	gaattacaag	caattcagct	agctttgcag	gattcagggc	4020
cagaagtaaa	catagtaaca	gactcacagt	atgcattagg	aatcattcaa	gcacaaccag	4080
ataagagtga	atcagaatta	gtcagtcaaa	taatagaaca	gttgataaaa	aaggaaaaag	4140
tctacctatc	atgggtacca	gcacataaag	gaattggagg	aaatgaacaa	gtagacaaat	4200
tagtaagtag	tggaatcaga	aaagtactgt	ttctagatgg	aatagataaa	gctcaagaag	4260
agcatgaaaa	atatcacagc	aattggagag	caatggctag	tgagtttaat	ctgccacca	4320
tagtagcaaa	ggaaatagta	gccagctgtg	ataaatgtca	gctaaaaggg	gaagccatgc	4380
atggacaagt	cgactgtagt	ccaggaatat	ggcaattaga	ctgtacacat	ttagaaggaa	4440
aatcatcct	agtagcagtc	catgtagcca	gtggctacat	ggaagcagag	gttatcccag	4500
cagaaacagg	acaagaaaca	gcatacttta	tactaaaatt	agcaggaaga	tggccagtca	4560
aagtaataca	tacagataat	ggcagtaatt	tcaccagtac	cgcagttaag	gcagcctgtt	4620
ggtgggcaga	tatccaacgg	gaatttggaa	ttccctacaa	tccccaaagt	caaggagtag	4680
tagaatccat	gaataaagaa	ttaaagaaaa	tcatagggca	agtaagagat	caagctgagc	4740
accttaagac	agcagtacaa	atggcagtat	tcattcacaa	ttttaaaaga	aaagggggga	4800
ttggggggta	cagtgcaggg	gagagaataa	tagacataat	agcatcagac	atacaaacta	4860
aagaattaca	aaaacaaatt	ataaaaattc	aaaattttcg	ggtttattac	agagacagca	4920
gagaccctat	ttggaaagga	ccagccaaac	tactctggaa	aggtgaaggg	gcagtagtaa	4980
tacaagataa	tagtgatata	aaggtagtac	caagaaggaa	agcaaaaatc	attaaggact	5040
atggaaaaca	gatggcaggt	gctgattgtg	tggcaggtag	acaggatgaa	gattagaaca	5100
tggcacagtt	tagtaaagca	ccatatgtat	gtttcgagga	gagctgatgg	atggttctac	5160
agacatcatt	atgaaagcag	acacccaaaa	gtaagttcag	aagtacacat	cccattagga	5220
gatgccaggt	tagtaataaa	aacatatttg	ggtctgcaga	caggagaaaag	agcttggcat	5280
ttgggtcacg	gagtctccat	agaatggaga	ttgagaagat	atagcacaca	agtagaccct	5340
gacctgacag	accaactaat	tcatatgcat	tattttgatt	gttttgcaga	atctgccata	5400
aggaaagcca	tactaggaca	gatagttagc	cctaagtgtg	actatcaagc	aggacataac	5460
aaggtaggat	ctctacaata	cttggcactg	acagcattga	taaaaccaa	aaagataaag	5520
ccacctctgc	ctagtgttag	gaaattagta	gaggatagat	ggaacaagcc	ccagaagacc	5580
aggggccgca	gaggaacca	tacaatgaat	ggacactaga	gcttttagaa	gaactcaagc	5640
aggaagctgt	cagacacttt	cctagaccat	ggctccataa	cttaggacaa	catatctatg	5700
aaacctatgg	agatacttgg	acaggagtgg	aagcaataat	aagaatcctg	caacaattac	5760
tgtttattca	tttcaggatt	gggtgccatc	atagcagaat	aggcattttg	cgacagagaa	5820
gagcaagaaa	tggagccaat	agatcctaac	ctagaaccct	ggaaccatcc	aggaagtacg	5880
cctaaaactg	cttgtaatgg	gtgttactgt	aaacgttgca	gctatcattg	tctagtttgc	5940
tttcagaaaa	aaggcttagg	catttactat	ggcaggaaga	agcggagaca	gcgacgaagc	6000
gctcctccaa	gcaataaaga	tcatcaagat	cctctaccaa	agcagtaagt	accgaatagt	6060
atatgtaatg	ttagatttaa	ctgcaagaat	agattctaga	ttaggaatag	gagcattgat	6120
agtagcacta	atcatagcaa	taatagtgtg	gaccatagta	tatatagaat	ataggaaatt	6180

ggtaaggcaa	aggaaaatag	actgggttagt	taaaaggatt	agggaaagag	cagaagacag	6240
tggcaatgag	agcgaggggg	atactgaaga	attatcgaca	ctgggtggata	tggggcatct	6300
taggcttttg	gatgctaattg	atgtgtaattg	tgaagggctt	gtgggtcaca	gtctactacg	6360
gggtacctgt	ggggagagaa	gcaaaaacta	ctctattttg	tgcatcagat	gctaaagcat	6420
atgagaaaga	agtgcataat	gtctgggcta	cacatgcctg	tgtacccaca	gaccccaacc	6480
cacaagaagt	gattttgggc	aatgtaacag	aaaattttta	catgtggaaa	aatgacatgg	6540
tggatcagat	gcaggaagat	ataatcagtt	tatgggatca	aagccttaag	ccatgtgtaa	6600
aattgacccc	actctgtgtc	acttttaaact	gtacaaatgc	aactgttaac	tacaataata	6660
cctctaaaga	catgaaaaat	tgctctttct	atgtaaccac	agaattaaga	gataagaaaa	6720
agaaagaaaa	tgactttttt	tatagacttg	atatagtacc	acttaataat	aggaagaatg	6780
ggaatattaa	caactataga	ttaataaatt	gtaatacctc	agccataaca	caagcctgtc	6840
caaaagctctc	gtttgaccca	attcctatac	attattgtgc	tccagctggg	tatgcgcctc	6900
taaaatgtaa	taataagaaa	ttcaatggaa	taggaccatg	cgataatgtc	agcacagtac	6960
aatgtacaca	tggaaattaag	ccagtgggat	caactcaatt	actgtttaat	ggtagcctag	7020
cagaagaaga	gataataatt	agatctgaaa	atctgacaaa	caatgtcaaa	acaataatag	7080
tacatcttaa	tgaatctata	gagattaaat	gtacaagacc	tggcaataat	acaagaaaga	7140
gtgtgagaat	aggaccagga	caagcattct	atgcaacagg	agacataata	ggagatataa	7200
gacaagcaca	ttgtaacatt	agtaaaaaatg	aatggaatac	aactttacaa	agggttaagtc	7260
aaaaattaca	agaactcttc	cctaataagta	cagggataaa	atttgacca	cactcaggag	7320
gggacctaga	aattactaca	catagcttta	attgtggagg	agaatttttc	tattgcaata	7380
caacagacct	gtttaatagt	acatacagta	atggtacatg	cactaatggg	acatgcatgt	7440
ctaataatac	agagcgcatc	acactccaat	gcagaataaa	acaaattata	aacatgtggc	7500
aggaggtagg	acgagcaatg	tatgccctc	ccattgcagg	aaacataaca	tgtagatcaa	7560
atattacagg	actactatta	acacgtgatg	gaggagataa	taatactgaa	acagagacat	7620
tcagacctgg	aggaggagac	atgagggaca	attggagaag	tgaattatat	aaatacaagg	7680
tggtagaaat	taaaccatta	ggagtagcac	ccactgctgc	aaaaaggaga	gtgggtggaga	7740
gagaaaaaag	agcagtagga	ataggagctg	tgctccttgg	gttcttggga	gcagcaggaa	7800
gcactatggg	cgagcatcac	ataacgctga	cggtacaggc	cagacaatta	ttgtctggta	7860
tagtcacaac	gcaaagtaat	ttgctgaggg	ctatagaggc	gcaacagcat	atgttgcaac	7920
tcacggctctg	gggccttaag	cagctccagg	caagagtcct	ggctatagag	agatacctac	7980
aggatcaaca	gctcctagga	ctgtggggct	gctctggaag	actcatctgc	accactaatg	8040
tgctttggaa	ctctagttag	agtaataaaa	ctcaaagtga	tatttgggat	aacatgacct	8100
ggatgcagtg	ggatagggaa	attagtaatt	acacaaacac	aatatacagg	ttgcttgaag	8160
actcgcaaag	ccagcaggaa	agaaatgaaa	aagattttact	agcattggac	agggtggaaca	8220
atctgtggaa	ttgggttagc	ataacaaatt	ggctgtggta	tataaaaaata	ttcataatga	8280
tagtaggagg	cttgataggt	ttaagaataa	tttttgctgt	gctctctcta	gtaaatagag	8340
ttaggcaggg	atactcacc	ttgtcattgc	agacccttat	cccaaaccgc	aggggacccg	8400
acaggctcgg	aggaatcgaa	gaagaagggtg	gagagcaaga	cagcagcaga	tccattcgat	8460
tagtgagcgg	attcttgaca	cttgccctggg	acgacctacg	aagcctgtgc	ctcttctgct	8520
accaccgatt	gagagacttc	atattaattg	tagtgagagc	agtggaaactt	ctgggacaca	8580
gtagtctcag	gggactgcag	aggggggtggg	gaacccttaa	gtatttgggg	agtcttgtgc	8640
aatattgggg	tctagagtta	aaaaagagtg	ctattaatct	gcttgatact	atagcaatag	8700
cagtagctga	aggaacagat	aggattctag	aattcataca	aaacctttgt	agaggatatcc	8760
gcaacgtacc	tagaagaata	agacagggct	tcgaagcagc	tttgcaataa	aatggggggc	8820
aagtggcaca	aaagcagtat	aattggatgg	cctgaagtaa	gagaaagaat	cagacgaact	8880
aggtcagcag	cagagggagt	aggatcagcg	tctcaagact	tagagaaaca	tggggcactt	8940
acaaccagca	acacagccca	caacaatgct	gcttgccctc	ggctggaagc	gcaagaggag	9000
gaaggagaag	taggctttcc	agtcagacct	caggtacctt	taagaccaat	gacttataaa	9060
gcagcaatag	atctcagctt	ctttttaaaa	gaaaaggggg	gactggaagg	gttaattttac	9120
tccaagaaaa	ggcaagagat	ccttgatttg	tgggtttata	acacacaagg	cttcttccct	9180
gattggcaaa	actacacacc	gggaccaggg	gtcagatttc	cactgacctt	tggatggtac	9240
ttcaagctag	agccagtcga	tccaagggaa	gtagaagagg	ccaatgaagg	agaaaacaac	9300
tgtttactac	accctatgag	ccagcatgga	atggaggatg	aagacagaga	agtattaaga	9360
tggaaagttg	acagtacgct	agcacgcaga	cacatggccc	gcgagctaca	tccggagtat	9420
tacaaagact	gctgacacag	aagggacttt	ccgctggggc	tttccactgg	ggcgttccag	9480
gaggtgtggg	ctgggcggga	caggggagtg	gtcagccctg	agatgctgca	tataagcagc	9540
tgcttttcgc	ctgtactggg	tctctctagg	tagaccagat	ctgagcccg	gagctctctg	9600

gctatctagg gaaccactg cttaagcctc aataaagctt gccttgagtg ccttgagtag 9660
tgtgtgcccg tctgttgtgt gactctggta actagagatc cctcagacca cttgtggtag 9720
tgtggaaaat ctctagca 9738

<210> 46

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: spacer

<400> 46

Gly Gly Gly Ser

1